

(11) EP 0 933 423 B1

(12)

EUROPEAN PATENT SPECIFICATION

- (45) Date of publication and mention of the grant of the patent: 22.08.2007 Bulletin 2007/34
- (21) Application number: 96935358.0
- (22) Date of filing: 17.10.1996

(51) Int Cl.: C12N 15/57 (2006.01) C12N 15/12 (2006.01) C07K 14/435 (2006.01) C07K 15/18 (2006.01)

C12N 15/12(2006.01) C07K 14/435 (2006.01) C07K 16/18 (2006.01) C12P 21/02 (2006.01) A61K 39/395 (2006.01) C12N 9/64 (2008.01) C07K 14/47 (2008.01) C12P 21/08 (2008.01) C12N 1/21 (2008.01) A61K 38/17 (2008.01) A61K 31/70 (2008.01)

- (86) International application number: PCT/JP1996/003017
- (87) International publication number:WO 1997/031109 (28.08.1997 Gazette 1997/37)

(54) MELTRINS

MELTRINE

MELTRINES

- (84) Designated Contracting States:

 AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC

 NL PT SE
- (30) Priority: 23.02.1996 JP 6175696
- (43) Date of publication of application: 04.08.1999 Bulletin 1999/31
- (83) Declaration under Rule 28(4) EPC (expert solution)
- (60) Divisional application: 07011948.2
- (73) Proprietor: MOCHIDA PHARMACEUTICAL CO., LTD. Shinjuku-ku, Tokyo 160-0004 (JP)
- (72) Inventors:
 - FUJISAWA, Atsuko, Room 305 Koganei-shi, Tokyo 184 (JP)
 - YAMAKAWA, Toru, Mochida Pharmaceutical Co., Ltd. Tokyo 160 (JP)
 - SHIRAKAWA, Kamon, Mochida Pharmaceutical Co., Ltd. Tokyo 160 (JP)

- MiZUSHIMA, Chitose, Mochida Pharmaceutical Co. Ltd Tokyo 160 (JP)
- OGAWA, Naoki, Mochida Pharmaceutical Co., Ltd. Tokyo 160 (JP)
- (74) Representative: Woods, Geoffrey Corlett J.A. KEMP & CO.
 Gray's Inn
 14 South Square
 London WC1R 5JJ (GB)
- (56) References cited:

WO-A-97/09420

WO-A-97/40072

- EMBL DATABASEAccession no D50411, Sequence Indentity MMMAB 6 November 1995 FUJISAWA SEHARA A: "Mouse mRNA for meltrin alpha" XP002151067
- EMBL DATABASEAccession no D50410, Sequence indentity MMMBA 6 November 1995 FUJISAWA SEHARA A: "Mouse mRNA for meltrin beta" XP002151068
- EMBL DATABASEAccession no D50412, Sequence indentity MMMGC 6 November 1995 FUJISAWA SEHARA A: "Mouse mRNA for meltrin gamma" XP002151069

EP 0 933 423 B1

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deerned to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

- WESKAMP G ET AL: "MDC9, A WIDELY EXPRESSED CELLULAR DISINTEGRIN CONTAINING CYTOPLASMIC SH3 LIGAND DOMAINS" THE JOURNAL OF CELL BIOLOGY, US,ROCKEFELLER UNIVERSITY PRESS, vol. 132, no. 4, 1 February 1996 (1996-02-01), pages 717-726, XP000644308 ISSN: 0021-9525
- NOMURA N ET AL: "PREDICTION OF THE CODING SEQUENCES OF UNIDENTIFIED **HUMAN GENES. I. THE CODING SEQUENCES OF** 40 NEW GENES (KIAA0001-KIAA0040) **DEDUCEDBY ANALYSIS OF RANDOMLY** SAMPLED CDNA CLONES FROM HUMAN IMMATURE MYELOID CELL LINE KG-1" DNA RESEARCH, JP, UNIVERSAL ACADEMY PRESS, vol. 1, no. 1, 1994, pages 27-35, XP002049267 ISSN: 1340-2838 - & EMBL DATABASE Accession no D14665, Sequence Identity HSORF0931 March 1993 NOMURA N ET AL: "Prediction of new human genes......" XP002151070 -& SWISSPROT **DATABASE Accession no Q13443 01 November** 1996 "Cellular disintegrin-related protein precursor* XP002151071
- CELL STRUCTURE AND FUNCTION, 20(6), (1995),
 T. YAGAMI-HIROMASA et al., "Meltrinalpha, A Novel Metalloprotease-Disintegrin, Participates in Myotube Formation", p. 585 2A-1345.

- CELL STRUCTURE AND FUNCTION, 20(6), (1995),
 T. KURISAKI et al., "Cloning of Meltrins, Novel Mambers of Metalloprotease-DisIntegrin Family Involved in Myogenesis", p. 585 2A-1330.
- NATURE, 377, (1995), T. YAGAMI-HIROMASA et al., "A Metalloprotease-Disintegrin Participating in Myoblast Fusion", pages 652-656.
- GENERALIZED RESEARCH REPORT IN FISCAL YEARS HEISEI 5 TO 7, MORPHOLOGICAL AND BIOCHEMICAL AND MOLECULAR BIOLOGICAL BASIC RESEARCH ON MUSCULAR DYSTROPHY, (03.96), JUNKO FUJISAWA et al., "Molecular Mechanism of Differentiation and Formation of Skeletal Muscle", pages 85-88.
- GENOMICS, 34(3), (15.06.96), C. CHO et al.,
 "Chromosomal Assignment of Four Test is Expressed Mouse Genes from a New Family of Transmembrane Proteins (ADAMs) involved in Cell-Cell Adhesion and Fusion", pages 413-417.
- EXPERIMENTAL MEDICINE, 14(10), (01.07.96), TOMOHIRO KURIZAKI, JUNKO FUJISAWA (SEBARA), "A Novel Cellular Adhesive Molecule Meltrin Belonging to Metalloprotease Dysintegrain Family", pages 1352-1356.
- BIOCHEMISTRY, 68(8), (08.96), TSUYOSHIENDO,
 "ADAM Family and Cell Fusion", pages 1453-1458
- PROC. NATL. ACAD. SCI. U.S.A., 91, (1994), G.
 WESKAMP et al., "A Family of Cellular Proteins Related to Snake Venom Disintegrins".

Description

Technical Field

[0001] This invention relates to Meltrins and polypeptides of the respective domains thereof; DNAs encoding the same; antisense oligonucleotides for these DNAs; various antibodies against these Meltrins and the polypeptides of the respective domains thereof; expression vectors comprising the DNAs; transformants constructed by using these expression vectors; a process for producing the above-mentioned meltrins and the polypeptides of the respective domains thereof by means of the transformants; and medical compositions comprising the Meltrins or Meltrin antagonists as an effective ingredient.

Background Art

[0002] In the course of myotube formation, myoblasts, which have divided from myogenic cells originating in undifferentiated mesodermal cells and grown to differentiate, will start synthesizing muscle-specific substances such as myosin and actin after its final division, and will lose cell boundaries at the fusion surface to be tansformed into multinucleate syncytium named myotube through adhesion and fusion of cytoplasmic membranes with neighbouring cells of the same kind.

[0003] There have been already reported several kinds of membrane proteins involved in the myotube formation, such as N-Cadherin (Knudsen, K.A. et al., Expl. Cell Res., 188, 175-184 (1990), Merge, R.M. et al., J. Cell Sci., 103, 897-906 (1992)), M-Cadherin (Donalies, M. et al., Proc. Natl. Acad. Sci., U.S.A. 88, 8024-8028 (1991)), N-CAMs (Merge, R.M. et al., J. Cell Sci., 103, 897-906 (1992) and others), V-CAMs and Integrins (Rosen, G.D. et al., Cell 69, 1107-1119 (1992) and others).

[0004] However, the molecular mechanism has not yet been sufficiently understood concerning the course of formation of the multinucleate syncytium named myotube through adhesion and fusion of the cytoplasmic membranes of the myoblasts with each other.

[0005] On the other hand, the substances named "fusion peptides" have been known as an adhesion factor involved in the course of infectior of cells with viruses (Morrison, T.G. Virus Res., 10, 113-136 (1988) and the others). Fertilin, which was recently isolated as a factor involved in sperm-egg adhesion, has been found to contain a sequence similar to the fusion peptide of rubella virus (Biobel, C.P. et al., Nature 356, 248-252 (1992) and the others).

[0006] Many substances having adhesion activity are known as mentioned above, and substances which may inhibit the activity of Integrins and the like have been developed and studied as potential medical agents.

[0007] The present inventors have now isolated novel substances involved in adhesion. Particularly, on the assumption that some fusion peptide-like adhesion factor like in sperm-egg adhesion may be involved in adhesion and fusion of the myoblasts with each other in the course of myotube formation, the novel substances involved in cell adhesion have been cloned and named "Meltrins", by using highly conserved sequences in Fertilin α and β as a probe.

Disclosure of Invention

[0008] The present invention relates to a novel "Meltrin." "Meltrins" are characterized as proteins which are expressed in the course of differentiation-induction of muscle cells and to contain the highly conserved sequences in Fertilin α and β. Meltrins are also characterized as proteins which are involved in fusion, adhesion, or aggregation of cells. Thus, some kinds of cells such as muscle ones may fuse, aggregate or adhere via Meltrins.

[0009] Cell fusion means that more than two cells fuse with each other to form one multicleate syncytium. Adhesion of cells means that more than two cells adhere to each other. Aggregation of cells means that more than two cells (particularly the cells present in liquid) flock together to form a mass of cells. It may be considered that cells adhere to each other, followed by cell fusion and aggregation.

[0010] According to the invention, there is thus provided a soluble meltrin polypeptide which does not comprise a transmembrane domain or an intracellular domain and which comprises the amino acid sequence of Gly (No.1) to Ile (No. 686) from the N-terminal in Fig. 15a - Fig. 15f, or the amino acid sequence of Glu (No. 156) to Ile (No. 686) from the N-terminal in Fig. 15a - Fig. 15f.

[0011] The invention also provides:

55

- a DNA comprising a base sequence encoding the polypeptide of the invention;
- a DNA of the invention which comprises the base sequence of No.1 to No. 2058 from the 5' terminal in Fig. 15a Fig. 15f;
 - a DNA of the invention which comprises the base sequence of No. 1 to No. 2848 from the 5' terminal in Fig. 15a Fig. 15f;

- an antisense oligonucleotide which hybridizes with a part of the sequence of No. 1957 to No. 2848 from the 5' terminal in Fig. 15a Fig. 15f;
- an antibody which recognizes the C-terminal region of a meltrin wherein the C-terminal region is from amino acid No. 653 to No. 686 from the amino terminal in Fig. 15a - Fig. 15f;
- a vector comprising a DNA of the invention;

25

- a transformant by the vector of the invention;
- a process for producing a polypeptide of the invention, which process comprises culturing the appropriate transformant of the invention:
- a medical composition comprising a polypeptide, an antisense oligonucleotide or an antibody of the invention;
- use of a polypeptide, an antisense oligonucleotide or an antibody of the invention for the manufacture of a medicament for treatment of a condition associated with unhealthy enhanced bone resorption; and
 - use of a polypeptide, an antisense oligonucleotide or an antibody of the invention for the manufacture of a medicament for preventing metastasis of cancer cells.
- At least three kinds of molecules (α, β) and (α, β) have been isolated from one animal species.
 - [0012] Meltrins may be mouse Meltrins α , β and γ , which are characterized by amino acid sequences shown in Fig. 2a Fig.2i, Fig.3a Fig.3j and Fig.4a Fig.4i, respectively, or partial sequences thereof.
 - [0013] Other examples are human Meltrins α , β and γ , which are characterized by amino acid sequences shown in any one of Fig.12a Fig.12b, Fig.15a Fig.15f or Fig.23a Fig.23b; any one of Fig.16 or Fig.17a Fig.17c; or Fig.13a Fig.13d, respectively, or partial sequences thereof.
 - [0014] The above amino acid sequences should be considered only examples of Meltrins. Any variant of the above amino acid sequences wherein a part of the sequences has changed due to deletion, substitution, addition, insertion and the like of amino acids is therefore a Meltrin, as long as it is expressed in muscle cells, and have the highly conserved sequences in Fertilin α and β or is involved in fusion, adhesion or aggregation of cells. As shown now by the present inventors, a high homology is seen in the part from disintegrin domain to cysteine-rich region of mouse amino acid sequences shown in Fig.2a Fig.2j and human amino acid sequences shown in Fig.12a Fig.12b. It is considered that such substances as showing homology of about 80 % or more, preferably about 90 % or more to the above amino acid sequences may keep the function as Meltrin. Particularly, it is believed that the substances having the sequences with homology of about 80 % or more, preferably about 90 % or more to the region from metalloproteinase domain to disintegrin domain of mouse or human Meltrins α , β and γ will have substantially the same activity, even if all of the other sequences are different from them. Accordingly, Meltrins may include substances having a high homology to the above amino acid sequences or to a part thereof and showing substantially the same activity as mouse or human Meltrins.
 - [0015] In other words, Meltrins may be characterized by having amino acid sequences encoded by base sequences that may hybridize the sequences complementary to the base sequences encoding any one of the amino acids shown in Fig.2a Fig.2j, Fig.3a Fig.3j, Fig.4a Fig.4i, Fig.12a Fig.12b, Fig.13a Fig.13d, Fig.15a Fig.15f, Fig.16, Fig.17a Fig.17c or Fig.23a Fig.23b.
 - [0016] Meltrins exist in bodies as a membrane protein consisting of intracellular domain, transmembrane domain, and extracellular domain; and as a soluble protein having no transmembrane domain. The extracellular domain contains precursor domain, metalloproteinase domain, disintegrin domain, and cysteine-rich region. Meltrin α has a fusion peptide-like sequence in its cysteine-rich region (Refer to Fig.8).
 - [0017] The disintegrin domain is indispensable for the function of Meltrins such as adhesion, fusion and aggregation of cells. On the other hand, the precursor and metalloproteinase domains are thought to be regulating sequences for Meltrins to show the activity in a specific organ or tissue, or under specific conditions. It is known that the disintegrin found in snake venom will adhere to platelet IIb/IIIa. It is therefore presumed that the disintegrin domain by itself may have the function to adhere to cells. The metalloproteinase domain may act by itself as a protease as such.
 - [0018] A polypeptide may comprise any part of a Meltrin. Such polypeptides include the respective domain per se of Meltrins, polypeptides comprising at least the respective domain of Meltrins, any part of the sequences of Meltrins, polypeptides comprising at least any part of the sequences of Meltrins, and polypeptides comprising at least the sequence having the combination of any of the respective domains of Meltrins and any part of Meltrins in any order.
- [0019] The above polypeptides which are chemically modified or formed into salts thereof.
 - [0020] The preferable examples of the such polypeptides include polypeptides consisting of a part of the disintegrin domain, polypeptides consisting of the disintegrin domain per se, polypeptides comprising at least the disintegrin domain, polypeprides comprising at least the disintegrin and cysteine-rich regions, polypeprides comprising at least the metalloproteinase, disintegrin and cysteine-rich regions, polypeptides consisting of the metalloproteinase domain, and polypeptides consisting of the metalloproteinase domain per se.
 - [0021] There may be mentioned as other preferable examples of the present polypeptides those comprising at least the disintegrin and cysteine-rich regions, but not comprising the transmembrane domain, or comprising neither the transmembrane domain nor intracellular domain; and those comprising at least the metalloproteinase, disintegrin and

cysteine-rich regions, but not comprising the transmembrane domain, or comprising neither the transmembrane domain nor intracellular domain. Such polypeptides comprising no transmembrane domain are a soluble one which will be secreted through a cell membrane into extracellular area. The soluble polypeptides may be collected from supernatant of the culture medium of cells. When optionally combined downstream of a suitable signal sequence and expressed by cells in a genetic engineering process, it will be secreted into the culture supernatant and advantageously collected therefrom with a high efficiency.

[0022] The amino acid sequences in Fig.2a - Fig.2j, Fig.3a - Fig.3j, Fig.4a - Fig.4i, Fig.12a - Fig. 12b, Fig.13a - Fig. 13d, Fig.15a - Fig.15f, Fig.16, Fig.17a - Fig.17c and Fig.23a - Fig.23b, which correspond to the precursor domain, metalloproteinase domain, disintegrin domain, cysteine-rich region, intracellular domain, and transmembrane domain of mouse and human Meltrins α , β and γ , are discussed in the Examples. It should be noted, however, that the polypeptides having the above corresponding amino acid sequences constitute only examples. That is to say, polypeptides essentially comprising the same amino acid sequences may also be prepared. Thus, the boundaries of each domain are not limited to those defined in the Examples. Polypeptides comprising the domains wherein the boundaries are shifted to N-, C-terminals or both by 1 to about 20 amino acids from the boundaries defined in the Examples may also be prepared, as long as they have substantially the same function as that of the above polypeptides. Similarly, the polypeptides wherein a part of the amino acid sequences has changed due to deletion, substitution, addition, insertion and the like of amino acids may also be prepared, as long as they have substantially the same function as that of each domain.

[0023] It is considered that the polypeptides comprising such amino acid sequences as showing homology of about 80 % or more, preferably about 90 % or more to the amino acid sequences in each domain of the above figures may have the same function as that of the polypeptide of the present invention.

[0024] The Meltrin of the present invention may be used to bond cells to each other or to apparatuses such as a plate. It may be also fused with any other substances to efficiently deliver the substances to muscle cells upon its application into culture systems of the muscle sells, tissues or bodies.

[0025] On the other hand, the polypeptides comprising at least a part of Meltrins may be added to the culture systems to competitively inhibit the adhesion, fusion or aggregation of cells. Particularly, the disintegrin domain per se, a part thereof, or a soluble polypeptide comprising the disintegrin domain may be used as an effective ingredient in a medical composition for inhibiting the adhesion of cells. For example, such medical composition may be used as an anticoagulant to inhibit thrombus formation or blood coagulation, and be used to treat thrombosis, DIC and multi-organ failure. Furthermore, since it is considered that adhesion factors such as integrin family are involved in metastasis of cancer cells, the polypeptides comprising the disintegrin domain may be used as a drug for inhibiting the growth of cancers, or the adhesion of cancer cells to other cells so as to prevent their metastasis. In addition to the above, it is known that the adhesion of cells plays an important role in formation of osteoclast. The examples will demonstrate that Meltrins are involved in the adhesion in the formation of osteoclast, and anti-Meltrin antibodies may inhibit the formation of osteoclast and the increase of bone resorption. Accordingly, the polypeptide of the present invention may be used as an effective ingredient in a medical composition for inhibiting the increase of bone resorption, like as anti-Meltrin antibodies,

[0026] Among the polypeptides comprising at least a part of the Meltrin of the present invention, those comprising the metalloproteinase domain may act as a protease by itself, or be used to competitively inhibit the activity of other proteases so that they may be utilized as a drug for treating inflammatory diseases.

[0027] The Meltrin polypeptide of the present invention may also be used as an antigen for producing antibodies.

[0028] The present invention also relates to DNAs comprising the base sequence encoding the amino acid sequences of the Meltrin of the present invention or the polypeptides comprising any parts thereof.

[0029] The above DNAs include any type of DNAs such as genomic DNAs and cDNAs.

35

[0030] Examples of DNAs encoding Meltrins are those encoding mouse Meltrins α , β , and γ , or the polypeptides comprising any parts thereof, which are characterized by the coding regions shown as the base sequences in Fig.5a - Fig.5j, Fig.6a - Fig.6h, and Fig.7a - Fig.7e, respectively, or partial sequences thereof. Other examples are those encoding human Meltims α , β , and γ , or the polypeptides comprising any parts thereof, which are characterized by the coding regions of the sequences shown as the base sequences in any one of Fig.12a - Fig.12b, Fig. 15a - Fig. 15f or Fig.23a - Fig.23b; any one of Fig.16 or Fig.17a - Fig.17c; or Fig.13a - Fig.13d, respectively, or partial sequences thereof.

[0031] The base sequences in the above figures, which correspond to the precursor domain, metalloproteinase domain, disintegrin domain, cystein-rich domain, intracellular domain, and transmembrane domain of mouse and human Mettrins α , β and γ , are discussed in the Examples. It should be noted, however, that they constitute only examples of such DNAs. DNAs essentially comprising the same base sequences may also be prepared.

[0032] Thus, the boundaries of each domain are not limited to those defined in the Examples. And the DNAs comprising sequences encoding the domains wherein the boundaries are shifted to 5'-, and/or 3'-ends by 1 to about 60 base pairs from the boundaries defined in the Examples may also be prepared, as long as they encode the polypeptides having substantially the same function as that of each domain.

[0033] In addition of the above base sequences, DNAs may be prepared comprising the base sequences or partial sequences thereof, which encode the same amino acid sequences as above prepared by means of chemical synthesis

or genetic engineering in consideration of degeneracy of codons. As now shown by the present inventors, a high homology is seen in mouse and human Meltrins. It is therefore considered that the substances showing homology of about 80 % or more, preferably about 90 % or more to the above amino acid sequences may keep the function as Meltrin, and that DNAs encoding such homologous polypeptides will hybridize with each other. Accordingly, DNA fragments may be obtained by hybridization under stringent conditions using the DNAs having the base sequences complementary to those in the above figures as a probe.

[0034] The DNAs of mouse or human Meltrins α , β and γ , or partial sequences thereof may be inserted into plasmid vectors. Strains of E. coli transformed by the same plasmid vectors have been deposited with the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology.

[0035] The DNAs described herein may be prepared by known methods. The cDNAs, for example, may be prepared by using cDNA library and known PCR (e.g., Michael A.I. et al., PCR Protocols, a guide to method and application, Academic Press, 1990) with degenerative primers for a part of the amino acid sequences (for example, the degenerative primer encoding the amino acid sequences of the disintegrin domain) shown in Fig.2a - Fig.2j, Fig.3a - Fig.3j, Fig.4a - Fig.4i, Fig.12a - Fig. 12b, Fig.13a - Fig.13d, Fig.15a - Fig.15f, Fig.16, Fig.17a - Fig.17c and Fig.23a - Fig.23b. The DNAs described herein may also be prepared by hybridization method using a probe prapared on the basis of the base sequences of the above amplified DNA fragments.

[0036] As demonstrated in the Examples, the preferable source of cDNA library include cells obtained by inducing myoblast to differentiate, bone marrow and fetal pulmonary cells. Known cDNA libraries prepared from placenta, chorionic cells and fetal cells may also serve as the source of cDNA library in the present invention.

[0037] Among the DNAs described herein, one encoding the polypeptide in which any parts of Meltrins are combined in any order may be prepared by the following steps. That is, each DNA fragment encoding any part of Meltrins is amplified by PCR, in which the primers may be optionally modified in order to provide an appropriate restriction enzyme site. The amplified DNA fragments are ligated with each other by DNA ligase, so that a reading frame should not be shifted. [0038] The DNAs described herein may be used for producing the Meltrins or polypeptides of the present invention by means of genetic engineering. Such prodution may be carried out with reference to known methods (for example, Sambrook J. et al., Molecular Cloning a Laboratory Manual 2nd ed., Cold Spring Harbor Laboratory, New York, 1989). [0039] The DNAs described herein inserted into suitable vectors may also be used in gene therapy. The base sequence encoding any physiologically active substances is fused downstream of the present DNAs followed by insertion of the resulting fused DNA into a vector originated in an appropriate virus, and cells in a living body are transformed with the resulting vector, so that the physiologically active substances may be expressed as a fused protein with the Meltrin of the present invention. The thus expressed physiologically active substances will be delievered near to the cells to which Meltrins adhere

[0040] The present invention further relates to antisense oligonucleotides and derivatives thereof for the DNAs encoding the Meltrin of the present invention or for the polypeptides comprising any part thereof.

[0041] The present antisense oligonucleotides and derivatives thereof are characterized by their base sequences complementary to those encoding Meltrins or a part thereof, or by their function to inhibit the expression of Meltrins or the polypeptides comprising any part thereof. The antisense oligonucleotides and derivatives thereof characterized by the latter feature include those complementarily bonding to the non-coding regions existing upstream or downstream of the coding regions of Meltrins as well as those complementarily bonding to the coding regions of Meltrins or any part thereof.

[0042] Examples of the antisense oligonucleotides and derivatives thereof described herein include the base sequences complementary to the DNAs of the present invention or any part thereof, particularly to those shown in Fig. 5a - Fig. 5j, Fig.6a - Fig.6h, Fig.7a - Fig.12a - Fig.12b, Fig.13a - Fig.13d, Fig. 15a - Fig. 15f, Fig.16, Fig.17a - Fig.17c and Fig.23a - Fig.23b. Uracil (U) may be used instead of thymine (T) as a complementary base to adenine (A).

[0043] The derivatives of the present antisense oligonucleotides include any one that is similar to the antisense oligonucleotides in steric structure and function, such as those wherein other substances are bound to 3'- or 5'-end of the oligonucleotides; those wherein at least one of bases, sugars or phosphoric acids in the oligonucleotides has substitution or modification; those having non-naturally occurring bases, sugars or phosphoric acids; and those having back bone other than that of sugars-phosphoric acids.

[0044] The antisense oligonucleotide of the invention and derivatives thereof may be prepared by known methods (for example, ed., Stanley T. Crooke and Bernald Lebleu, in Antisense Research and Applications, CRC Publishing, Florida, 1993).

[0045] The present antisense oligonucleotide of a naturally occurring type may be prepared by chemically synthesizing sense-primers and antisense-primsers having the base sequences complementary to 3'- or 5'-end of the antisense oligonucleotide sequences, followed by PCR using the Meltrin genes or RNAs encoding Meltrins as a template. Otherwise, the derivatives of the antisense oligonucleotides such as a methylphosphonate and phosphorothionate types may be prepared by means of a chemical synthesizer (e.g., Perkin Elmer Japan Co., Type 394) according to the manual attached to the chemical synthesizer, followed by, if necessary, purification of the synthesized products in HPLC method using

reversed phase chromatography and the like.

25

[0046] The present antisense oligonucleotide and derivatives thereof may be labelled with radioisotopes, fluorescent substances, enzymes or luminescent substances and used as a probe for detecting the existence of Meltrins or any part thereof in a sample. The present antisense oligonucleotide may also be used as a medical composition for inhibiting the expression of Meltrins in a living body.

[0047] For the purpose of inhibiting the expression of Meltrins by using the present antisense oligonucleotide and derivatives, they may be solubilized or suspended in a suitable solvent, enclosed in a liposome, or inserted into a suitable vector.

[0048] It is preferred that the present antisense oligonucleotide and derivatives thereof used in the medical composition should have a pharmaceutically acceptable purity and be used in a pharmaceutically acceptable way.

[0049] As already mentioned in the above, it is considered that Meltrins are involved in formation of osteoclast, growth and metastasis of cancers as well as skeletal myogenesis. Accordingly, the present antisense oligonucleotide and their derivatives which are capable of inhibiting the expression of Meltrins may be used in in the manufacture of a medicament treatment and prevention of cancers, treatment of osteoporosis and hypercalcemia by inhibiting bone resorption.

[0050] The present invention also relates to antibodies recognizing the Meltrin of the present invention or the polypeptides comprising at least any part thereof. In other words, they include those recognizing only Meltrins of the present invention, those recognizing only the polypeptides of the present invention and those recognizing both of them.

[0051] Antibodies described herein include those cross reacting with other polypeptides in addition to those specifically recognizing Meltrins and the polypeptides of the present invention. They also include those specifically recognizing any one of Meltrins α , β and γ , and those specifically recognizing more than two of Meltrins α , β and γ , as well as those recognizing only Meltrins originated in a particular animal such as human and mouse or only the polypeptides comprising at least any part thereof, and those recognizing Meltrins originated in more than two kinds of animals or the polypeptides comprising at least any part thereof.

[0052] Such antibodies include those recognizing the amino acid sequences in Fig.2a - Fig.2j, Fig.3a - Fig.3j, Fig.4a - Fig.4i, Fig.12a - Fig.12b, Fig.13a - Fig.13d, Fig.15a - Fig.15f, Fig.16, Fig.17a - Fig.17c or Fig.23a - Fig.23b, or any part thereof.

[0053] More preferably, the antibodies described herein are those obtained by immunization of animals with the polypeptides comprising said amino acid sequences or any part thereof as an antigen, which may be optionally conjugated with a suitable carrier.

[0054] Such antibodies may be prepared by inserting DNA comprising the base sequnces shown in Fig.5a - Fig.5j, Fig.6a - Fig.6h, Fig.7a - Fig.7e, Fig.12a - Fig.12b, Fig.13a - Fig.13d, Fig.15a - Fig.15f, Fig.16, Fig.17a - Fig.17c or Fig. 23a - Fig.23b or any part thereof into a suitable expression vector, tranforming a suitable host cell by the vector to produce Meltrins, which are purified from cell bodies of the transformant or culture medium and administered as an antigen. The cell bodies per se of the transformant or any cells expressing Meltrins per se may be administered as an antigen. Such transformant or cells may express any one of Meltrins α , β and γ , or more than two kinds of them. The antibodies may be also prepared by chemically synthesizing the polypeptides having a part of the amino acid sequences of Meltrins, conjugating them with a carrier such as KLH (Keyhole Limpet Hemocyanin) and administering them as an antigen.

[0055] It is possible to prepare the present antibody that may recognize the whole of Meltrins even when the part of Meltrins is used as an antigen to be administered. It is also possible to prepare the present antibody that may recognize human Meltrins or any part thereof even when mouse Meltrins or any part thereof are used as an antigen to administered. [0056] The antibodies described herein include monoclonal and polyclonal ones, and may belong to any class or subclass.

[0057] The antibodies may be prepared according to known methods (e.g., "Meneki jikkenho (Laboratory manual of Immunology)" published by Japan Immunological Society). An example of the known methods will be described below. [0058] A suitable cell is transformed by an expression vector comprising the coding regions of the base sequences shown in Fig.5a - Fig.5j, Fig.6a - Fig.6h, Fig.7a - Fig.7e, Fig.12a - Fig.12b, Fig.13a - Fig.13d, Fig.15a - Fig.15f, Fig.16, Fig.17a - Fig.17c or Fig.23a - Fig.23b or any part thereof, and used as an antigen as such. Alternatively, Meltrins produced by the transformant are purified from cell bodies of the transformant or culture medium to be used as an antigen, or polypeptides consisting of amino acid sequences shown in the above figures are chemically synthesized, cojugated with a carrier such as KLH (Keyhole Limpet Hemocyanin) and purified to be used as an antigen.

[0059] Animals are inoculated with the antigen thus prepared, alone or together with a suitable adjuvant such as Freund's complete adjuvant (FCA) or Freund's incomplete adjuvant (FIA), subjected to boosting at two to four-week intervals. After boosting, the blood is drawn from the animals and antiserum is obtained therefrom. Animals to be immunized may be selected from rat, mouse, rabbit, sheep, horse, fowl, goat, pig, cattle and the like, depending on the kind of the antibody to be desired. Polyclonal antibodies may be obtained by purification of the antiserum by known methods such as salting-out, ion-exchange chromatography, affinity chromatography and any combination thereof.

[0060] Monoclonal antibodies may be prepared as follows. Antibody-producing cells such as spleen cells and lym-

phocytes are collected from the immunized animals, fused with myeloma and the like by known methods using polyethyleneglycol, Sendai virus, electrical pulse to give hybridomas. Clones which produce an antibody bonding to the Meltrin of the present invention are then selected and cultured. Monoclonal antibodies of the present invention are purified from the culture supermatant of the selected clones by known methods such as salting-out, ion-exchange chromatography, affinity chromatography and any combination thereof.

[0061] The present antibodies may be neutralizing antibodies, which inhibit the fusion, adhesion or aggregation of cells by Meltrins. The neutralizing antibodies of the present invention include those that can completely inhibit the activity of Meltrins, and those partially inhibit the same.

[0062] The neutralizing antibodies may be screened by adding antiserum or culture supernatant of the hybridomas to the culture system of Meltrin-expressing cells to evaluate the degree of inhibition of fusion or aggregation of cells. After the screening, the desired antibodies may be purified from the thus selected antiserum or culture supernatant of the hybridomas by the known methods.

[0063] The antibodies of the present invention include Fab, F(ab'), F(ab'), and Fv, as long as they recognize and bond to the present polypeptides or any part thereof. A single chain Fv may be also included in the present antibodies, which is obtained by constructing a gene encoding the single chain Fv wherein H and L chains are linked into a single chain and being expressed by a suitable host cell. Chimera antibodies, human antibodies and humanized antibodies are also included in the present invention, as long as they recognize and bond to the present polypeptides or any part thereof. [0064] For example, the chimera antibodies may be prepared by substituting a gene encoding the constant region of human antibodies for a gene encoding the constant region of the mouse antibodies recognizing Meltrins or the polypeptides of the present invention, expressing the thus reconstituted gene in animal cells. The human antibodies may be prepared by, for example, in vitro sensitization method (Borrebaeck, C.A.K.J. Immunol., Meth., 123, 157, 1989) or the method using SCID mouse (Toshio KUDO, Tissue Culture, 19, 61-65, 1993). The humanized antibodies may be prepared by reconstituting a gene so that complementary determining regions (CDR) of the human antibodies are replaced with those of the mouse antibodies, and expressing the gene in animal cells (Carter et al., Pro. Nat. Acad. Sci, 89, 4285, 1992). [0065] If necessary, amino acids in a framework of the variable region of the humanized antibodies thus reconstituted may be replaced, so that the framework should have a high homology to that of the mouse antibodies and CDR of said humanized antibodies may form an appropriate antigen-binding site. The preferred examples of the humanized antibodies are those having the same CDR as the neutralizing antibodies F932-15-2 and F937-9-2. For the preparation of these preferred humanized antibodies, the DNA encoding the antibodies is prepared from the hyndoma F932-15-2 or F937-9-2, and linked with the DNAs encoding human antibodies so that the sequences other than CDRs should originate in the human antibodies. Any variation may be optionally introduced into the DNA encoding the framework portion. The thus obtained DNA is then inserted into a suitable expression vector to transform a suitable cell, and the humanized antibodies are purified from the culture supernatant of the transformant.

[0066] The present antibodies may be labelled with fluorescent substances, enzymes, luminescent substances or radioisotopes to detect Meltrins or their decomposed products present in body fluid or tissues. Since it is considered that Meltrins are involved in formation of myotube, resorption of bone and metastasis of cancers as already mentioned in the above, the detection of the existence of Meltrins in body fluid or tissues would make it possible to estimate the progress of diseases and prognosis and to confirm the effects of treatments. The present antibodies may be also used to provide an antibody affinity column, or to detect Meltrins in a fraction during the course of purification of Meltrins.

[0067] The neutralizing antibodies of the present invention may serve as an effective ingredient of a medical composition for inhibiting bone resorption, inflammatory diseases, blood coagulation and metastasis of cancers, owing to their ability to inhibit fusion or adhesion of cells. They may serve as an agent used in culture to inhibit the aggregation of cultured cells. When used as the effective ingredient of the medical composition, the human or humanized antibodies are preferred from the viewpoint of their antigenicity.

[0068] Also, the present invention relates to a vector comprising the DNA of the present invention. The present vector may further contain, if necessary, an enhancer sequence, promoter sequence, ribosome-binding sequence, base sequence for amplification of the number of copies, sequence encoding signal peptides, sequences encoding other polypeptides, poly(A)-additional sequence, splicing sequence, origin of replication, base sequence of the gene for selective markers and so on.

[0069] The present vector may be prepared by inserting the DNAs of the present invention into any vectors according to known methods (e.g., Molecular Cloning, a Laboratory Manual 2nd ed., Cold Spring Harbor Laboratory, New York, 1989). The preferable examples of the DNAs encoding Meltrins or any part thereof have been already disclosed in the present specification. The present vectors include a plasmid vector, phage vector and virus vector; pUC118, pBR322, pSV2-dhfr, pBluescriptll, PHIL-S1, λZap II, λgt10, pAc700, YRP17, pEF-BOS and pEFN-II being preferred.

[0070] The preferred vectors of the present invention may optionally comprise the origin of replication, selective markers, and promoter in addition to the DNAs encoding Meltrins or the polypeptides comprising at least any part thereof so as to be used to express Meltrins or the same polypeptides. As the origin of replication, ColEl, R factor, F factor and so on may be used in the vectors for E.coli; SV40- or adenovirus-derived ones in the vectors for animal cells; and ARS1-

derived one in the vectors for yeast. As the promoter, trp, lac and tac promoters may be used in the vectors for E. coli; SV40-, cytomegalovirus-, and adenovirus-derived ones, and those intrinsically existing in the genes of human or animals such as the promoter region of an elongation factor 1α in the vectors for animal cells; and α promoter in the vectors for yeast, especially AOX1 promoter in the case of Pichia yeast. In the addition to the above sequences, the present vectors may further comprise, if necessary, RNA splicing site, signal for poly-adenylation and the like for the transforamtion of eucaryotic cells. The present vectors may be used for the production of Meltrins or any part thereof by means of genetic engineering, and used in gene therapy for Meltrins-related diseases.

[0071] The present invention therefore relates to transformants transformed by the above vectors.

[0072] The present transformants may be prepared by transforming suitable host cells by the above vectors according to known methods (e.g., Idenshi Kogaku Handbook (Handbook of gene technology), extra edition of Jikkenigaku, Yodo, 1991)). The host cells may be selected from procaryotic ones such as E.coli and Bacillus, or eucaryotic cells such as yeast, insect cells, and animal ones. The preferred transformants of the present invention are those derived from E.coli, yeast or CHO cell as a host cell to express Meltrins or the polypeptides of the present invention.

[0073] The present invention further relates to a process for producing Meltrins or the present polypeptides comprising at least any part thereof, comprising the step of culturing the above transformants.

[0074] In the present producing process, the transformants of the present invention are cultured, optionally with amplification of the gene or expression-induction, if necessary, according to known methods (e.g., Biseibutsugaku Jikkenho (Laboratory manual of microbiology), Tokyo Kagaku Dojin, 1992). The culture mixture, i.e., the cells and culture supernatant, is collected and optionally subjected to concentration, solubilization, dialysis, and various chromatography to purify Meltrins or the present polypeptides comprising any part thereof. The purification of the present polypeptides may be carried out by an optional combination of the above known methods for the purification of proteins, and an efficient purification could be performed by using an affinity column with the antibodies of the present invention.

[0075] In the present producing process, the polypeptides of the present invention may be produced by the transformants as a fused protein with other proteins such as β -galactosidase. In such case, the fused protein should be treated with chemicals such as cyanogen bromide or enzymes such as protease in a certain step in the purification process, so that the polypeptides of the present invention may be excised.

[0076] The present invention relates to medical compositions comprising a novel effective ingredient, which is Meltrins of the present invention or Meltrin-antagonist. The "Meltrin-antagonist" means a molecule which is able to inhibit fusion, adhesion or aggregation of cells through Meltrins. It includes, for example, the present antibodies recognizing Meltrins and having a neutralizing activity, the fragments of the same antibodies, the polypeptides consisting of any part of Meltrins or any combination thereof in any order, the antisense oligonucleotides for the DNAs encoding Meltrins or derivatives thereof.

[0077] The antibodies recognizing Meltrins may be prepared by the methods already mentioned in the above, and from which the antibodies which may completely or partially neutralize fusion, adhesion or aggregation of muscle cells, osteoclast or cancer cells are selected and used as the effective ingredient of the present medical composistions. The antibodies to be used as the effective ingredient include those prepared by administering any polypeptides as the antigen into any animals, as long as they may recognize human Meltrins and inhibit fusion, adhesion or aggregation of human muscle cells, osteoclast or cancer cells. They may be polyclonal or monoclonal ones, being preferably the human or humanized antibodies, considering the fact that the medical compositions will be administered to human. The human or humanized antibodies may be prepared according to the methods already described in the above.

35

[0078] The above fragments to be used as the effective ingredient in the present medical compositions include Fab, F(ab'), F(ab'), and F(ab'), and F(ab') and F(ab')

[0079] The polypeptides having any part of Meltrins or any combination thereof in any order may be used as the effective ingredient of the medical compositions, as long as they have the activity of inhibiting fusion, adhesion or aggregation of cells.

[0080] The preferable examples of the above polypeptides include those comprising a part or the whole of the disintegrin domain of Meltrins, those comprising the metalloproteinase, disintegrin and cysteine-rich regions of Meltrins, those comprising the disintegrin domain, but not comprising the transmembrane domain of Meltrins, and those comprising at least the metalloproteinase and disintegrin domains, but not comprising the transmembrane domain of Meltrins. These polypeptides may be chemically synthesized or produced by means of genetic engineering, as already mentioned in the above.

[0081] The antisense oligonucleotides or derivatives thereof to be used as the effective ingredient of the medical compositions may have any base sequences or any structure, as long as they are suitable for administration to human, and will complementarily bond to the gene for Meltrins to completely or partially inhibit their expression.

[0082] As already mentioned, Meltrins are involved in formation of osteoclast and metastasis of cancer cells. Accordingly, the medical comosition comprising the Meltrin-antagonist as the effective ingredient may be used for the purpose of inhibition of bone resorption or metastasis of cancers. The antagonist against human Meltrin α or β is more preferably used as the effective ingredient in the medical composition for inhibition of bone resorption, while the antagonist against

human Meltrin γ is more preferably used as the effective ingredient in the medical composition for inhibition of cancer metastasis.

[0083] The Meltrins or Meltrin antagonist used as the effective ingredient in the present medical composition may be formed into their salts or be modified with pharmaceutically acceptable chemical agents, as long as they will never lose their essential activities. There may be exemplified as the salts those with inorganic acids such as hydrochloric acid, phosphoric acid, hydrobromic acid and sulfuric acid; those with organic acids such as maleic acid, succinic acid, malic acid and tartaric acid.

[0084] The medical compositions of the present invention include those administered by any route such as oral, subcutaneous, intravenous, intramuscular, intraperitoneal, intracutaneous, and intraintestinal ones.

[0085] Any administration methods and intervals may be adopted. The present medical comopsitions may comprise depending on the administration route pharmaceutically acceptable auxiliaries such as fillers, packing agents, thickeners, binding agents, humidifying agents, disintegrating agents, surfactants, solution aids, buffers, pain-easing agents, preservatives and stabilizers. In the case of injections, for example, they may comprise stabilizers such as gelatin, human serum albumin (HSA) and polyethylene glycol; alcohols and saccharides such as D-mannitol, D-sorbitol, and glucose; and surfactants such as Polysorbate 80 (TM).

[0086] The medical compositions of the present invention may be mainly used for the prevention and treatment of osteoporosis and hypercalcemia, or the prevention of infiltration and metastasis of cancers.

[0087] The present medical compositions may be administered in an amount of about 0.1 - 100 mg/kg/day, preferably of about 1 - 50 mg/kg/day, more preferably of about 1 - 10 mg/kg/day, depending on the conditions or ages of patients, or administration routes. It may also be continuously administered by an intravenous drip, or administered by a single dose or doses at appropriate intervals per day.

[0088] The present medical compositions may be formulated according to the conventional manners. The injection, for example, may be formulated by dissolving the Meltrins or their antagonists aseptically prepared to a pharmaceutically acceptable purity into physiological saline, buffers and the like, followed by addition of gelatin or HSA, if necessary. Such injections may also be lyophilized, which will be dissolved into distilled water for the injections, physiological saline and the like when they are used.

[0089] The screening of the substances which may bind to Meltrins, inhibit the activity of Meltrins or regulate their expression may be carried out by using the Meltrins, various polypeptides, DNAs encoding them and the like.

30 Brief Description of Drawing

[0090]

20

25

35

40

Fig.1a - Fig.1b show the comparison between parts of mouse Meltrins α , β , γ (referred to as "M α ", "M β ", "M γ ") and the known sequences (macrophage specific antigen (MS2), Jararhagin (JR), fertilin- α (f α).

Fig.2a - Fig.2j show the amino acid sequence of mouse Meltrin α and its corresponding DNA sequence.

Fig.3a - Fig.3j show the amino acid sequence of mouse Meltrin β and its corresponding DNA sequence, wherein "N" means unidentified base.

Fig.4a - Fig.4i show the amino acid sequence of mouse Meltrin γ and its corresponding DNA sequence. "N" means unidentified base.

Fig.5a - Fig.5j show the result of DNA sequence analysis of the DNA inserted into pBSMel α , which comprises the base sequence encoding mouse Meltrin α . "N", "M", "W" and "S" mean unidentified bases.

Fig.6a - Fig.6h show the result of DNA sequence analysis of the DNA inserted into pBSMelβ, which comprises the base sequence encoding mouse Meltrin β. "N", "M", "W" and "S" mean unidentified bases.

Fig.7a - Fig.7e show the result of DNA sequence analysis of the DNA inserted into pBSMelγ, which comprises the base sequence encoding mouse Meltrin γ. "N", "M", "W" and "S" mean unidentified bases.

Fig.8 shows schematically the structures of Meltrins α , β , γ , δ MP, δ Pro.

Fig. 9 is a photograph of electrophoresis showing the result of Western blotting.

Fig. 10 is a photograph of electrophoresis showing the result of Northern blotting.

50 Fig.11a- Fig.11b show fusion-promoting activity of Meltrins for myoblast.

Fig.12a - Fig.12b show the result of base sequence analysis of the DNA inserted into pBShuMa300, which encodes human Meltrin α. "N" and "X" mean unidentified bases and unidentified amino acids, respectively.

Fig. 13a - Fig. 13d show the result of base sequence analysis of the DNA inserted into pBShuM γ G238, which encodes human Meltrin γ .

Fig. 14a shows schematically the cloning region in the cloning of human Meltrin α .

Fig.14b shows schematically the cloning region in the cloning of human Meltrin β .

Fig.15a - Fig.15f show partial amino acid sequence and its corresponding base sequence of human Meltrin α , determined based on the result of analysis of the DNA inserted into pMel α -26N, pMel α -25C.

Fig.16 shows amino acid sequence and its corresponding base sequence of human Meltrin β .

Fig.17a - Fig.17c show partial amino acid sequence and its corresponding base sequence of human Meltrin β , determined based on the result of analysis of the DNA inserted into pMel β -24C, pMel β -24N.

Fig.18a shows schematically the sites of the peptides administered as the antigens in mouse Meltrin α .

Fig.18b shows amino acid sequences of the peptides administered as the antigens.

Fig. 19 is a photograph of electrophoresis showing the result of Western blotting with anti-mouse Meltrin α antibodies. Fig. 20 is a graph showing the inhibition of myotube formation by anti-mouse Meltrin antibodies.

Pig.21 is a graph showing the effects by anti-mouse Meltrin antibodies on the formation of pit (bone-resorption area) by mouse all bone cells.

Fig.22 is a graph showing the effects on the serum Ca values of the mouse fed with low Ca-content feed by antimouse Meltrin antibodies.

Fig.23a - Fig.23b show the amino acid sequence comprising the transmembrane domain of human Meltrin α and its corresponding base sequence.

Fig.24a - Fig.24e show the result of base sequence analysis of the DNA inserted into pMelβ-24C, pMelβ-24N.

Best Mode for Carrying Out the Invention

[0091] The present invention will be further illustrated by the following Examples, which should not be construed to limit the scope of the present invention.

Examples

5

10

15

20

[0092] The abbreviations used in the following description are based on the conventional ones in the art.

[0093] The processes used in the following Examples are based on Sambrook J. et al., Molecular Cloning, a Laboratory Manual 2nd ed., Cold Spring Harbor Laboratory, New York, 1989; E. Harlow, D.Lane et al., Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory; and the like.

Example 1. Acquisition of the DNAs encoding mouse Meltrins by RT-PCR

30 (1) Preparation of RNA, cDNA.

[0094] Myogenic cell line derived from fetal fibroblast C3H10T1/2, (a clone transfected by the gene encoding "myogenin", a muscle differentiation-controlling factor and expressing the myogenin) was proliferated to the extent of 10⁶ cells/\$\phi\$ 10 cm plate in DMEM supplemented with 10% fetal bovine serum (MOREGATE) and cultured at 37°C for 2 days in differentiation medium (DMEM containing 2 % horse serum from GIBCO) for differentiation and induction. Total RNA was separated according to Guanidine isothiocyanate/acid phenol method (Chomczynski P. and Sacchi N., Anal. Biochem., 162, 156-159, 1987), and poly (A) RNA was selectively separated by repeating twice oligo(dT)-cellulose column chromatography. By using the poly(A) RNA as a template and random primers (N6, Pharmacia), cDNAs were synthesized with MLV reverse transcriptase (GIBCO BRL) according to its manual for synthesis. The obtained cDNAs were then used as a template for the next PCR, and double strand DNAs were synthesized and inserted into a phage (\$\pi\$ZapII (stratagene)) to give a cDNA library.

(2) RT-PCR

45 [0095] RT-PCR was carried out by using the cDNAs prepared in the above (1) as a template in the following steps: [0096] A degenerative primer encoding the amino acid sequence EDCDCG or EECDCG was synthesized and used as a sense primer, and a degenerative primer encoding the amino acid sequence KCGKLIC was synthesized and used as an antisense primer.

[0097] The primers were mixed with the above cDNAs, Taqpolymerase and the reaction agents (Boehringer Manheim), and subjected to 36 reaction cycles of 95°C for 1 min, 55°C for 2 min, and 72°C for 3 min. The amplification product of around 450 bp was then collected by 1.5% agarose gel electrophoresis.

[0098] The amplified fragments thus obtained were inserted into a Smal site in the plasmid pBS-SKII(-) (stratagene), and subjected to DNA sequence analysis by means of a DNA sequencer (370A type, Applied Biosystems). As a result, it was found that three kinds of molecules (DNA fragments) existed (Fig.1), which were then used as a probe to screen the cDNA library so as to isolate cDNAs comprising an open reading frame with 903, 920 and 845 amino acid residues, respectively (Fig.2a - Fig.2j, Fig.3a - Fig.3j, Fig.4a - Fig.4i). The products of the respective genes were named Meltrins α , β , and γ (Fig.5a - Fig.5j, Fig.6a - Fig.6h, Fig.7a - Fig.7e). These cDNAs were inserted into pBS-SKII(-) to give the plasmids, "pBSMel α ", "pBSMel β ", and "pBSMel γ ", respectively.

[0099] E.coli strain JM109 was transformed according to a known method by the above plasmids "pBSMelα", "pBS-Melβ", and "pBSMelα", respectively, and the resulting transformants "JM109(pBSMelα)", "JM109(pBSMelβ)", and "JM109 (pBSMelβ)" were deposited in the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken 350 Japan) on February 19, 1996 under accession numbers FERM P-15451, FERM P-15452, and FERM P-15453, respectively, and then transferred on October 8, 1996 to the deposit under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and Regulation under accession numbers FERM BP-5701, FERM BP-5702, and FERM BP-5703, respectively.

(3) Analysis of the structure of Meltrins

20

40

[0100] From the structure analysis of Meltrins on the basis of the DNA sequences determined in the above (2), it was supposed that Meltrins α , β , and γ were a transmembrane-type protein consisted of an extracellular domain, transmembrane (TM) domain, and intracellular domain, and that the extracellular domain consists of a precursor domain (pro region) comprising a signal peptide-like sequence, metalloproteinase domain, disintegrin domain, and the following cysteine-rich region. A fusion peptide-like sequence was contained in the cysteine-rich domain of Meltrin α (Fig.8).

[0101] Based on their homology to the snake venom, Jararhagin, it has been considered that in Meltrin α , the precursor domain corresponded to the sequence from N-terminal to Arg (No.205) and to the bases No.221-835, the metalloproteinase domain to the sequence from Glu (No.206) to Pro (No.414) and to the bases No.836-1462, the disintegrin domain to the sequence from Phe (No.420) to Gly (No.509) and to the bases No.1478-1747, the cysteine-rich region to the sequence from His (No.510) to Gly (No.706) and to the bases No.1748-2338, the fusion peptide-like sequence to the sequence from Gly (No.585) to Glu (No.607) and to the bases No.1973-2041, the transmembrane domain to the sequence from Leu (No.707) to Leu (No.727) and to the bases No.2339-2401.

[0102] Similarly, it was considered that in Meltrin β, the precursor domain corresponded to the sequence from N-terminal to Arg (No.204) and to the bases No.63-674, the metalloproteinase domain to the sequence from Glu (No.205) to Pro (No.409) and to the bases No.675-1289, the disintegrin domain to the sequence from Tyr (No.415) to Gly (No. 504) and to the bases No.1305-1574, the cysteine-rich region to the sequence from Thr (No.505) to Pro (No.706) and to the bases No.1575-2180, the transmembrane domain to the sequence from Val (No.707) to Arg (No.729) or to Leu (No.724) and to the bases No.2181-2249 or 2181-2234.

[0103] Similarly, it was considered that in Meltrin γ , the precursor domain corresponded to the sequence from N-terminal to Arg (No.205) and to the bases No.69-683, the metalloproteinase domain to the sequence from Ala (No.206) to Pro (No.406) and to the bases No.684-1292, the disintegrin domain to the sequence from Tyr (No.412) to Gly (No. 502) and to the bases No.1302-1574, the cysteine-rich region to the sequence from Tyr (No.503) to Ala (No.694) and to the bases No.1575-2150, the transmembrane domain to the sequence from Leu (No.695) to Ile (No.714) and to the bases No.2151-2210.

Example 2: Establishment of anti-Meltrin α antibodies

(1) Preparation of immunogen

[0104] A chimera polypeptide was prepared as follows, which consisted of glutathione-S-transferase (GST) (Smith, D.B. & Johnson, K.S., Gene, Vol.67, 31-40, 1988) and the polypeptide having the amino acid sequence from Ser (No. 483) to Lys (No.635) of Meltrin α in Fig.2a - Fig.2j, said polypeptide being attached to the C-terminal of GST. First, the plasmid, pGEX2T (Pharmacia) comprising the cDNA encoding GST was digested at a BamHI site and used as a vector. On the other hand, the cDNA corresponding to the amino acid sequence from Ser (No.483) to Lys (No.635) of Meltrin α in Fig.2a - Fig.2j was amplified from pBSMel α by PCR, and ligated with a BamHI linker by a DNA ligase. The resulting cDNA was then ligated with the above vector by a DNA ligase to give a plasmid, which was then tranformed into E.coli strain NM522.

[0105] The transformed E.coli was cultured in L-broth with 1mM IPTG to produce a large amount of the chimera polypeptide in the inclusion bodies upon expression-induction. The strain was suspended into MTPBS (150mM NaCl, 16mM Na₂HPO₄, 4mM NaH₂PO₄, 0.1mM PMSF), subjected to ultrasonication, and solubilized with 1% Triton. The supernatant of the thus treated mixture was collected. Glutathione agarose (Sigma) was mixed with the supernatant to adsorb the chimera polypeptide which was then eluted with an elution buffer (50mM Tris-HCl, pH 8.0, 0.5mM glutathione) and used as an immunogen.

(2) Preparation of antiserum

[0106] The antigen (1mg) prepared in the above (1) in 0.5ml PBS and RIBI in PBS 0.5ml (MPL+TDM+CWS Emulsion,

Funakoshi) was mixed with each other, and subcutaneously or intracutaneously administered into a rabbit (12 weeks old, female). After boosting three times with 500µg dose at 4 week intervals, the blood was collected and serum was separated to give antiserum.

5 (3) Affinity purification of antiserum

[0107] The chimera polypeptide expressed in E.coli and solubilized in the above (1), or GST having no fused polypeptide was bound to the glutathione agarose beads. The resulting beads were washed with 0.2M sodium borate (pH 9.0), and mixed with dimethyl pimelidiate (a final concentration of 20mM) so that the antigen was irreversibly bound to the beads, so as to give chimera polypeptide-affinity beads and GST-affinity beads, respectively.

[0108] The antiserum diluted ten times with 10mM Tris-HCl (pH 7.5) was first mixed with the GST-affinity beads for anti-GST antibodies to be absorbed and removed, and then mixed with chimera polypeptide-affinity beads for anti-Meltrin α antibodies to be adsorbed thereon. The resulting chimera polypeptide-affinity beads were washed with 10mM Tris (pH 7.5) and 500mM NaCl, and the anti-Meltrin α antibodies were eluted with 100mM glycine and collected as purified anti-Meltrin α antibodies.

(4) Western blotting

[0109] C2 cell was proliferated to the extent of 10⁶ cells/\$\phi\$ 10cm plate in DMEM supplemented with 15% fetal bovine serum, then cultured at 37°C in differentiation medium (DMEM supplemented with 2% horse serum) and collected on the second day (C2DM d2) and on the 4th day (C2DM d4).

[0110] Further, C2 cell transformed by pBOSMelα (+) prepared in the following Example 5 (3) was cultured in DMEM supplemented with 15% fetal bovine serum at 31°C for three days, inoculated into a plastic dish (φ 6cm) at a density of 2 x 10⁵/dish, further cultured for one day and transferred into the above differentiation medium for differentiation induction. After two day-culture in the differentiation medium, the cells were collected.

[0111] The collected C2DM d2, C2DM d4 or transformants by pBOSMela (+) was mixed with SDS solubilizing buffer (100mM Tris-HCI (pH 6.8), 4% SDS, 20% Glycerol), subjected to ultrasonication and centrifuged to give their supernatant as a sample.

[0112] A membrane was wahsed twice with a washing solution. The antiserum prepared in the above (3) was diluted 20 times with 5% skim milk solution in TBS-T, into which the membrane was soaked and incubated at 37°C for one hour. After the incubation, the membrane was washed twice with the washing solution. The membrane was then soaked into a biotin-labelled anti-rabbit immunoglobulin antibody (Daco) diluted 4,000 times with the above skim milk solution and incubated at 37°C for one hour. After the incubation, the membrane was washed twice with the washing solution. The membrane was reacted with a peroxidase-labelled streptoavidin for one hour, washed twice, stained with MB reagent (Cat.TM912, Shic) and detected by ECL system (Amersham).

[0113] The results are shown in Fig.9.

[0114] The Western blotting revealed the bands at about 115KD, 86KD, 67KD, and 58KD, indicating that Meltrin α was expressed as a glycoprotein. It was also considered that the precursor domain was deleted in the molecule of 86KD, and both the precursor and metalloproteinase domains were deleted in the molecule of 67KD or 56KD.

Example 3: Northern blotting

40

[0115] Poly (A)+ RNAs were prepared from vaious tissues of mouse (bone, brain, liver, heart and skeletal muscle of adult mouse; bone and skeletal muscle of newborn mouse; and bone and skeletal muscle of fetal mouse) by using a mRNA purification kit of Pharmacia according to the method described in Example 1. RNAs were denatured by heating at 65°C for 5 min in 50% formamide, subjected to electrophoresis on 1.5% agarose gel comprising 6.6% formalin, and transferred onto a nylon membrane (Highbond-N, Amersham).

[0116] On the other hand, cDNAs encoding a part of the disintegrin and cysteine-rich regions (Glu(No.434) - Cys(No. 583) in Fig.2a - Fig.2j, Glu(No.429) - Cys(No.578) in Fig.3a - Fig.3j, Glu(No.426) - Cys(No.575) in Fig.4a - Fig.4i) were prepared by PCR, and labelled with ³²P using a random primer labelling kit (Megaprime, Amersham). As a control probe, cDNA encoding G3PDH (glyceraldehyde 3-phosphate dehydrogenase) was also llabelled with ³²P in the same way. The above mRNAs were hybridized with the radiolabelled cDNAs under high stringency conditions according to the method of Sambrook J.et al. (Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Habor Laboratory, New York, 1989).

55 [0117] Their results are shown in Fig. 10.

[0118] Fig.10 has revealed that Meltrin α and β were expressed only in bones of adult and newborn mouses, and skeletal muscles of newborn and fetal mouses (the results from the fetal mouse are not shown in Fgi.10). There was no tissue-specificity in the expression of Meltrin γ , since it was universally expressed in all the tissues.

Example 4: Confirmation of adhering activity of Meltrin α

(1) Construction of plasmids pBOSMelαδMP(+) and pBOSMelαδMP(-)

[0119] A deletion type Meltrin δ MP wherein the precursor and metalloproteinase domains in the extracellual domain of Meltrin α had been deleted was prepared in the following method.

[0120] The plasmid, pBSMel α was partially digested at MscI and subjected to electrophresis on 1% agarose gel to give a linear plasmid DNA. The resulting DNA was partially digested at Nhel, treated with a Klenow fragment to generate blunt ends, and subjected to intramolecular ligation. Vectors having the right deletion were selected and their DNA sequences were confirmed. After digestion at multicloning sites of EcoRV and Notl in the vectors, a deletion type δ MP fragment of about 5.8kb was obtained.

[0121] On the other hand, the plasmid, pEFBOS (Mizushima S. & Nagata S, Nucleic Acid Res. Vol.18, p.5322, 1990) was digested by a restriction enzyme Xbal, dephosphorylated, treated with a Klenow fragment to generate blunt ends and subjected to electrophresis on 1% agarose gel to give a linear plasmid DNA. The resulting linear DNA was then ligated with the above fragment of about 5.8kb by a DNA ligase to give the plasmids pBOSMelαδMP(+) and pBOSMelαδMP(-). They were the constructs comprising the inserted DNA encoding the δMP fragment wherein the amino acid sequence of from lle(55) to Glu(399) of Meltrin α was deleted, in sense direction and antisense direction, respectively.

(2) Construction of plasmid pBOSMelα(+)

20

25

55

[0122] The plasmid, pBSMelα, was partially digested by EcoRV and Notl to give a fragment of about 7kb. The above pEFBOS plasmid was digested by a restriction enzyme Xbal, dephosphorylated, treated with a Klenow fragment to generate blunt ends, and subjected to electrophresis on 1% agarose gel to give a linear plasmid DNA. The resulting linear DNA was then ligated with the above fragment of about 7kb by a DNA ligase to give the plasmids pBOSMelα(+).

(3) Preparation of plasmid pBOSMelαδPro(+)

[0123] There was a AfIII site in the boundary region between the precursor and metalloproteinase domains of Meltrin α , and there was a Nhel site in the boundary region between metalloproteinase and disintegrin domains of Meltrin α . On the other hand, there remined the Nhel site in the boundary region between the signal peptide-like sequence and disintegrin domain in pBOSMel α 8MP(+) prepared in the above (1). Accordingly, pBOSMel α was digested at AfIII, ligated with a Nhel linker immediately before its metalloproteinase domain and digested at Nhel, so that the metalloproteinase would be excised. The excised domain was inserted into the Nhel site between the signal peptide-like sequence and the disintegrin domain of pBOSMel α 8MP(+) to give the expression plasmid, pBOSMel α 8Pro(+) encoding 8Pro wherein there a deletion was found around the precursor domain (the amino acid sequence of from Ile(No.55) to Glu(No.206) of Meltrin α).

(4) Confirmation of myoblast fusion-promoting activity

[0124] Myoblast cell line C2 was transfected by the mixture comprising the plasmid pBOSMelα(+) or pBOSMelαδMP (+), and the plasmid pSV2NEO in a molar ratio of 20:1 by using LIPOFECTAMINE (Gibco BRL) according to its protocol. The transfected cells were diluted and inoculated on a plate (φ 10cm) coated with collagen (IWAKI) so that the transformants would be obtained at a density of 10 - 20 clones per plate. The inoculated cells were cultured for 12 days in DMEM containing 20 % fetal bovine serum and 5 ng/ml of bFGF (Gibco BRL) followed by isolation thereof.

[0125] For the purpose of the examination of myoblast fusion-promoting activity, the resulting transformants and the parent strain C2 were cultured for 3-4 days in the absence of bFGF, inoculated onto a plastic dish (φ 6cm) at a density of 2 x 10⁵/dish, and further cultured for one day, followed by the 4 day culture in the above differentiation medium for differentiation induction. Upon differentiation induction, C2 began to form myotube. After the 4 day culture followed by fixation with methnol and staining with Giemsa and Wright's reagents (Merck), the number of nuclei were determined at any four independent fields of 1 mm² on the dish and fusion index was calculated as follows:

Fusion Index = 100 * (The number of nuclei in multicleate syncytium having three or more nuclei) / (The number of the total nuclei)

[0126] Further, the time course of the fusion index was observed after differentiation induction every one day for five days.

[0127] The results are shown in Fig.11a - Fig.11b. As seen from these figures, the fusion activity of the transformant expressing the full length of Meltrin α (pBOSMel α (+) which was referred to as "full length" in Fig.11a) become lower than that of the parent cell, and it was therefore considered that the full length of Meltrin α would suppress the cell fusion in some way. On the other hand, the transformant harboring pBOSMel α 8MP(+), which was referred to as " Δ MP" in the figures, significantly promoted the cell fusion activity. It was also observed that the transformant harboring pBOSMel α 8Pro (+) promoted the cell fusion activity.

[0128] On the other hand, the C2 cell transformed by the plasmid pBOSMel β (+) prepared by the insertion of the DNA encoding the full length of Meltrin β in the same way as in the above (2) could not cause any significant change in the fusion activity for muscle cells. However, The C2 transformant cotransfected by pBOSMel α (+) and pBOSMel β (+) promoted the cell fusion activity compared with that of parent cell.

[0129] On the other hand, neither the C2 cell transformed by the plasmid pBOSMel γ (+) prepared by the insertion of the DNA encoding the full length of Meltrin γ in the same way as in the above (2), nor the C2 transformant cotransfected by pBOSMel α (+) and pBOSMel γ (+) could cause any significant change in the fusion activity for muscle cells.

[0130] These results demonstrate that Meltrin α is involved in the fusion of muscle cells, and will show its activity to promote the cell fusion upon its processing. It is estimated that Meltrin α or Meltrin β does not act alone, but act in the form of a heteromer between them, since the transformant expressing both Meltrin α and Meltrin β promoted the fusion of muscle cells.

(5) Examination of the function of Meltrins in non-muscle cells

[0131] The mouse fibroblast L929 was transformed by pBOSMel α (+) or pBOSMel β (+) and the transformants expressing Meltrin α or Meltrin β were isolated. These transformants did not aggregate, nor fuse with each other. This was also true for the case of the transformant expressing both Meltrin α and Meltrin β .

[0132] On the othe hand, the L929 cells transformed by pBOSMely(+) could showed a significant aggregation activity upon the addition of calcium ion, after the cells had been torn from a plate in a medium comprising no calcium ion.

[0133] These results demonstrate that Meltrin γ has a cell aggregation activity, and by considering the similarity of these molucules it is suggested that myoblast fusion-promoting activity of Meltrin α and Meltrin β may be attributed to their myoblast aggregation-promoting activity.

Example 5: Inhibition of adhering activity by antisense

20

[0134] The plasmid BOSMelαδMP(-) prepared in Example 4 (1) was mixed with the plasmid PSV2NEO at a molar ratio of 20:1, by which C2 cells were tranformed according to the method of Example 4 (4) followed by isolation of the transformants expressing antisense RNA. The adhering activity of the thus isolated transformants was determined by the method of Example 4. The results are shown in Fig.11a - Fig.11b, which demonstrated that the fusion of C2 cells was inhibited by the expression of antisense RNA for δMP (referred to as "AS" in the figures).

[0135] The above results have revealed that Meltrin α plays an essential role in the cell fusion of muscle cells.

Example 6: Preparation of cDNA fragments encoding human Meltrins α and γ

[0136] By using mRNA purified from human myelocytes (Clonetech Co.) as a template, cDNAs were prepared according to the method of Example 1 (1), and 36 cycles of PCR was then carried out by using the degenerative primer obtained in Example 1 (2) and said cDNAs as a template. The amplified product was inserted into a EcoRV site of pBS-SKII(-), and named "pBShuMα300." The results of DNA sequencing are shown in Fig.12a and Fig.12b.

[0137] It was found that the DNA sequence comprised the base sequence encoding the part from an intermediate position of the disintegrin domain to an intermediate position of the cysteine-rich region of human Meltrin α (the disintegrin domain is located to Gly (No.36), followed by the cysteine-rich region in Fig.12a and Fig.12b).

[0138] On the other hand, by using a part of a human sequence (D-14665) registered with a data base, whose function had not yet identified, a senseprimer (5'-CACGATGATGGGAGAGATTG-3') and antisense primer (3'-CACTCTGATT-TCCTATGCCTC-5') were synthesized. PCR was carried out according to the above method to give the amplified product, which was then inserted into the EcoRV site of pBS-SKII(-), and named "pBShuMyG238." The results of DNA sequencing are shown in Fig.13a and Fig.13b.

[0139] It was found that the DNA sequence comprised the base sequence encoding the part from an intermediate position of the metalloproteinase domain to an intermediate position of the cysteine-rich region of human Meltrin γ (the metalloproteinase domain is located from N-termial to Pro (No.40), the disintegrin domain from Lys (No.41) to Gly (No. 136) or from Tyr (No.46) to Gly (No.136), followed by the cysteine-rich region from Tyr (No.137)). The E. coli strain JM109 was transformed by those plasmids to give JM109(pBShuMα300) and JM109(pBShuMγG238), which were deposited in the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology

(1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken 350 Japan) on February 19, 1996 under accession numbers FERM P-15454 and 15455, respectively, and then transferred on October 8, 1996 to the deposit under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and Regulation under accession numbers FERM BP-5704 and 5705, respectively.

Example7: Preparation of cDNA fragment encoding human Meltrin α by usig cDNA library derived from human placenta -1

(1) First screening

5

10

20

25

40

[0140] Based on the cDNA sequence of Meltrin α obtained in Example 6, sense primer MA-1 and antisense primer MA-2 were synthersized (see Table 1). The human placenta λ gt11 cDNA library (Clonetech Co., code No. CLHL1008b) was inoculated onto LB plate (ϕ 10cm) at such a density that 10,000 plaques per plate may be obtained. After the formation of plaques, SM buffer 5ml was added to each plate, the plates were put by incubation at a room temperature for 4 hours, and phages were collected from each plate (plate lysate method). PCR was carried out by using the collected phage solution as a template. Thus, mA-1 and MA-2 primers, Ex Taq polymerase (TaKaRa Co.,), and its reagents (TaKaRa Co.,) were mixed, followed by 35 cycles of the reactions at 94°C for 30sec, 55°C for 30sec, and 72°C for one min. A part of the amplified products was subjected to an agarose gel electrophoresis, and a phage solution of the clone comprising Meltrin α cDNA was selected.

(2) Second screening

[0141] The phage solution of the desired clone obtained in the first screening was inoculated at such a density that 400 plaques per plate may be obtained. After the formation of plaques, phages were collected in the same manner as above and a phage solution comprising the desired clone was selected.

- (3) Third screening
- [0142] The phage solution of the desired clone obtained in the second screening was inoculated at such a density that 40 plaques per plate may be obtained. After the formation of plaques, phages were collected in the same manner as above and a phage solution comprising the desired clone was selected.
 - (4) Forth screening
- [0143] The phage solution of the desired clone obtained in the third screening was inoculated at such a density that 10 plaques per plate may be obtained. After the formation of plaques, phages were collected in the same manner as above and a phage solution comprising the desired clone was selected.
 - (5) Final screening

[0144] The phage solution of the desired clone obtained in the forth screening was inoculated at such a density that 20 plaques per plate may be obtained. After the formation of plaques, each plaque was stuck with a toothpick, and the sticking material was suspended as a template into PCR solution. The above 35 cycles of the PCR with MA-1 and MA-2 primers finally gave two psitive clones. A single positive plaque comprising the desired clone was collected in SM buffer, and the phage was lysed thereinto.

[0145] PCR was carried out by using λ gt11 Forward primer and λ gt11 Reverse primer (Table 1) to give a fragment of human Meltrin α cDNA in the phage vector.

[0146] From a partial DNA sequencing of the terminal bases of the resulting fragments it was estimated that those cDNAs comprised the base sequences encoding human Meltrin α obtained in Example 6, and corresponded to about 650 amino acids (Clone 23) or about 500 amino acids (Clone 25) of mouse Meltrin (Fig. 14).

Example 8: Preparation of cDNA fragment encoding human Meltrin α by usig cDNA library derived from human placenta -2

[0147] A sense primer Mel α-5'S was designed based on the sequence encoding the N-terminal of the cDNA sequence of the clone 23 revealed in Example 7. The human placenta λgt11 cDNA library (Clonetech Co.) was screened by the sense primer Mel α-5'S and antisense primer MA-2 to give cDNA encoding about 700 amino acids (Clone 26) (Fig.14a). For the purpose of the analysis of the base sequence of Meltrin gene, the four primers, λgt11 Forward-Eco, λgt11

Reverse-Eco, MA-1-Eco, and MA-2-Eco were synthesized (Table 1).

5

10

15

50

[TABLE 1] The base sequences of the primers for PCR

MA-1	: 5' ACG ATG GGC ACT CAT GTC AG 3'
MA-2	: 5' CAT CTC GCA TTT GGC AAA GG 3'
λ gt11 Forward	: 5' GGT GGC GAC GAC TCC TGG AGC CCG 3'
λ gt11 Reverse	: 5' TTG ACA CCA GAC CAA CTG GTA ATG 3'
Mel α-5' <i>S</i>	: 5' CAC TGA ACA TTC GGA TCG TG 3'
λ gt11 Forward-Eco	:5'CCGGAATTCGGTGGCGACGACTCCTGGAGCCCG3'
λ gt11 Reverse-Eco	: 5' CCG GAA TTC TTG ACA CCA GAC CAA CTG GTA ATG 3'
MA-1-Eco	: 5' CCG GAA TTC ACG ATG GGC ACT CAT GTC AG 3'
MA-2-Eco	: 5' CCG GAA TTC CAT CTC GCA TTT GGC AAA GG 3'
S-hMel α-TM5'	: 5' GCA CAA AGT GTG CAG ATG GA
A-mMel α-3'	: 5' CAG AGG CTT CTG AGG AGG N

[0148] The second half of the Meltrin gene was amplified by PCR using Clone 25 as a template, and MA-1-Eco and λ gt11 Reverse-Eco primers. The first half of the Meltrin gene was amplified by PCR using Clone 26 as a template, and MA-2-Eco and λ gt11 Forward-Eco primers. These cDNA fragments were digested at EcoRI and cloned into the EcoRI site of pUC 118 to give the plasmid vectors "pMel α -26N" and "pMel α -25C", respectively. The sequences of Meltrin α cDNA comprised in these plasmids were determined by a conventional method.

[0149] The E.coli strain JM109 was transformed by those plasmids according to the known method of Hanahan et al. to give JM109(pMelα-26N) and JM109 (pMelα-25C), and were deposited in the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken 350 Japan) on October 3, 1996 under the terms of the Budapest Treaty on the International Recognition of the Deposit of microorganisms for the Purposes of Patent Procedure and Regulation under accession numbers FERM BP-5689 and 5688, respectively.

[0150] The base sequence and its corresponding amino acid sequence of human Meltrin α which had been revealed by the base sequencing of pMel α -26N and pMel α -25C are shown in Fig.15a - Fig.15f.

[0151] Comparison of the DNA sequence thus obtained with that obtained in Example 6 indicated four discrepancies in base pairs, the three of which being silent mutation, and the other dicrepancy causing substitution of Asp (No.505) in the above figures for Glu in the sequence of Example 6.

[0152] The analysis of the structure of the base sequence showed that the DNA encoded the sequence from an intermediate part of the precursor domain to the C-termial of Meltrin α . Thus, it has been considered that in the amino acid sequence shown in Fig.15a - Fig.15f, the partial sequence (C-terminal) of the precursor domain corresponds to the the sequence from Gly N-terminal to Arg (No.155) and to the bases No.1-465, the metalloproteinase domain to the sequence from Glu (No.156) to Pro (No.364) and to the bases No.466-1092, the disintegrin domain to the sequence from Glu (No.365) or Phe (No.370) to Gly (No.459) and to the bases No.1093 or 1108-1377, the cysteine-rich region to the sequence from His (No.460) to Gln (No.656) or Ala (No.652) and to the bases No.1378-1968 or 1956, the fusion peptide-like sequence to the sequence from Gly (No.535) to Gln (No.557) and to the bases No.1603-1671. There was no transmembrane domain in this sequence, suggesting that human Meltrin α existed as a soluble protein without a transmembrane domain in a body. In other words, it is considered that Meltrin α having the amino acid sequence of Fig. 15a - Fig.15f is extracellularly secreted and present in blood or body fluid. It is considered that such soluble Meltrin α takes a part in regulating adhesion, fusion and aggregation of cells in the body.

[0153] It is considered that Meltrin α having the amino acid sequence of Fig.15a - Fig.15f has generated as a result of an alternative splicing of the gene. It is also considered that the DNA encoding the region downstream of the cysteinerich region, and the DNA encoding transmembrane doamin and intracellular domain are located on different exons, and that the splicing out of either DNA would yield a soluble type Meltrin, or a membrane-binding type Meltrin.

Example 9: preparation of cDNA fragments encoding human Meltrins β

- (1) Preparation of cDNA fragment encoding a part of the disintegrin domain of human Meltrin β
- 65 [0154] By using mRNA purified from human myelocytes (Clonetech Co.) as a template, cDNAs were prepared according to the method of Example 1 (1), and 36 cycles of PCR were then carried out by using the degenerative primers obtained in Example 1 (2) and said cDNAs as a template. The amplified product was inserted into pBS-SKII(-). The analysis of the resulting DNA sequence revealed that it was a partial sequence of Meltrin β. The determined DNA

sequence is shown in Fig.16.

- (2) First screening by using cDNA library originated in human fetal lung
- 5 [0155] Based on the partial cDNA sequence of Meltrin β obtained in the above (1), sense primer MA-3 and antisense primer MA-4 were synthersized (see Table 2). The human fetal lung λgt11 cDNA library (Clonetech Co., code No. CLHL1072) was inoculated onto LB plate (φ 10cm) at such a density that 10,000 plaques per plate may be obtained. After the formation of plaques, SM buffer 5ml was added to each plate. And the plates were put at a room temperature for 4 hours, and phages were collected from each plate (plate lysate method). PCR was carried out by using the collected phage solution as a template. Thus, MA-3 and MA-4 primers, Ex Taq polymerase (TaKaRa Co.,), and its reagents (TaKaRa Co.,) were mixed, followed by 35 cycles of the reactions at 94°C for 30sec, 55°C for 30sec, and 72°C for one min by means of DNA thermal cycler (Perkin Elmer Co.,). A part of the amplified products was subjected to an agarose gel electrophoresis, and a phage solution of the clone comprising Meltrin β cDNA was selected.
- 15 (3) Second screening
 - [0156] The phage solution of the desired clone obtained in the first screening was inoculated at such a density that 1000 plaques per plate may be obtained. After the formation of plaques, phages were collected in the same manner as above and a phage solution comprising the desired clone was selected.
 - (4) Third screening

20

25

40

- [0157] The phage solution of the desired clone obtained in the second screening was inoculated at such a density that 100 plaques per plate may be obtained. After the formation of plaques, phages were collected in the same manner as above and a phage solution comprising the desired clone was selected.
 - (5) Forth screening
- [0158] The phage solution of the desired clone obtained in the third screening was inoculated at such a density that
 10 plaques per plate may be obtained. After the formation of plaques, phages were collected in the same manner as
 above and a phage solution comprising the desired clone was selected.
 - (6) Collection and confirmation of DNA fragment comprising partial cDNA sequence
- 35 [0159] The PCR was carried out using the phage solution of the desired clone obtained in the forth screening (#24) as a template, and a combination of λgt11 Forward primer (Table 1) and MA-4 primer or a combination of λgt11 Reverse primer (Table 1) and MA-3 primer to give amplified products with about 500bp (24-F/4) and about 5kbp (24-R/3), respectively. From a partial DNA sequencing of the terminal bases of the resulting two DNA fragments, it was estimated that those cDNA comprised the base sequences determined in the above (1).
 - (7) Analysis of base sequences
 - [0160] For the purpose of subcloning of the cDNA fragments comprising the cDNA partial sequence of human Meltrin β, two primers MA-3-Eco and MA-4-Eco were newly synthersized (see Table 2).
- [0161] The PCR was carried out using the phage solution (#24) as a template, and a combination of λgt11 Forward-Eco primer (Table 1) and MA-4-Eco primer or a combination of λgt11 Reverse-Eco primer (Table 1) and MA-3-Eco primer. The resulting amplified products were digested with EcoRI and inserted into the EcoRI site of pUC118 to give the plasmids, "pMeIβ-24C" and "pMeIβ-24N", respectively. The sequence of Meltrin β cDNA comprised in these plasmids was determined by a conventional method.
- 50 [0162] The E. coli strain JM109 was transformed by those plasmids according to the known method of Hanahan et al. to give JM109(pMelβ-24C) and JM109 (pMelβ-24N), and were deposited in the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken 350 Japan) on October 3, 1996 under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and Regulation under accession numbers FERM BP-5690 and 5691, respectively.
 - [0163] The base sequence and its corresponding amino acid sequence which had been revealed by the base sequencing of pMelβ-24C and pMelβ-24N are shown in Fig.24a Fig.24e.
 - [0164] Comparison of the DNA sequence thus obtained with that obtained in the above (1) showed one discrepancy

in base pairs, which was a silent mutation, causing no change of amino acid.

[0165] The analysis of the structure of the base sequence showed that the DNA encoded the sequence from an intermediate part of the metalloproteinase domain to the C-termial of human Meltrin β . Thus, it has been considered that in the sequence shown in Fig.24a - Fig.24e, the partial sequence at C-terminal of the metalloproteinase domain corresponds to the the sequence from Gly (N-terminal) to Pro (No.36) and to the bases No.2-109, the disintegrin domain to the sequence from Asp (No.37) or Tyr (No.42) to Gly (No.131) and to the bases No.110 or 125-394, the cysteine-rich region to the sequence from Thr (No.132) to Pro (No.330) and to the bases No.395-991, the transmembrane domain to the sequence from Val (No.331) to Met (No.348) or Arg (No.353) and to the bases No.992-1045 or 1060. It is considered that the sequence from Tyr (No.349) or Gln (No.354) corresponds to the intracellular domain. However, as homology analysis to mouse Meltrin β shows a very low homology in the sequence from Pro (No.395), it is estimated that the sequence up to His (No.394) is involved in the function of extracellular domain of human Meltrin β . The sequence up to Pro (No.395) in Fig. 24a - Fig.24e is shown in Fig.17a - Fig.17c.

[TABLE 2] The base sequences of the primers for PCR

MA-3	: 5' TGC TGC CAC CAG TGT AAG 3'
IVIA-3	. 5 TGC TGC CAC CAG TGT AAG 3
MA-4	: 5' TCC TGG TAG GTG AGG CAC ATG 3'
MA-3-Eco	: 5' CCG GAA TTC TGC TGC CAC CAG TGT AAG 3'
MA-4-Eco	:5'CCG GAATTCTCCTGGTAG GTG AGG CAC ATG 3'

Example 10: Preparation of anti-Meltrin α monoclonal antibodies

(1) Selection of peptides

15

20

25

35

45

[0166] Based on the amino acid sequence of mouse Meltrin α determined in Exmple 1, their epitopes were analysed. [0167] Eight kinds of peptide sequences were selected as a potential epitope, based on the secondary structure estimated from the regions wherein discrepancy in amino acids is seen between Meltrins α and β , the estimated non-RGD region, and the region wherein metalloproteinase had been cleaved (Fig.18a and b). These eight kinds of peptides were synthesized by Peptide Synthesizer (ABI 432A) so that they would have Cys at their C-terminal, cleaved, and purified by HPLC of a reverse phase column (YMC-ODS).

(2) Preparation of antiserum

[0168] After lyophilization of the peptides obtained in the above (1), each peptide 0.55mg was dissolved in 0.1 M phosphate buffer (pH 7.0) 55µl. Maleimidated KLH (Boehringer Manheim) 0.77mg was dissolved in distilled water 77µl. The two resulting solutions were combined, and reacted at a room temperature for two hours, followed by the purification by Nick column (Pharmacia) equilibrated with physiological saline to give antigens to be used in the following experiments. [0169] Each antigen 50µg was diluted with physiological saline to 0.1 ml, mixed with the same amount of Freund's complete adjuvant (DIFCO) and administered intraperitoneally into Wistar rat (5 weeks old, female). The antigen was mixed with the same amount of Freund's incomplete adjuvant (DIFCO) and administered two weeks later in the same way as above.

(3) Evaluation of antiserum (plate assay)

[0170] After one week from the administration, the blood was drawn from the eyeground of the rat, and an increase of the antibody titer for the administered peptides was confirmed by the reaction between immobilized peptides and the antiserum according to a plate assay as follows.

[0171] First, 50mM phosphate buffered saline (0.9% NaCl, pH 7.2) comprising 0.5mg/ml of Sulfo-SMCC (Pierce) was poured into each well of an amino plate (Sumitomo Bakelite). After incubation at 37°C for 2 hours, the wells were washed five times with ion-exchanged water, and the above buffer comprising 0.5μg/ml of each peptide was added. After incubation at 37°C for one hour, the well were blocked by 0.076M phosphate buffered saline (0.45% NaCl, pH 6.4), which will be referred to hereinafter as "PBS", comprising 0.1% of BSA and 4mg/ml of cysteamine. The blocking agent was removed, each antiserum diluted 1,000 to 100,000 times with PBS was added followed by incubation at 37°C for one hour. After two repeats of washing of the wells with 0.9% NaCl comprising 0.005% Tween20, an anti-rat immunoglobulin abtibody labelled with peroxidase (Dako) and diluted with PBS comprising 10% rabbit serum was added to each well followed by incubation at 37°C for one hour. Upon the completion of the reaction, the wells were washed five times with

a washing liquid and two times with ion-exchanged water. And 0.1M McIlvaine buffer (pH 5.0) comprising 3mg/ml of ophenylene diamine and 0.027% hydro peroxide was added and reacted for 5min. The reaction was terminated by the addition of 1N HCl, and absorbance at 490nm was measured. The results are shown in Table 3, in which (++) means a strong reactivity, and (+) means a week reactivity.

TABLE 3 Reaction of antiserum with the peptide antigens

peptide antigens	Reaction of Antiserum
1 ProA	++
2 MP-A	++
3 MP-B	++ .
4 DC-A	· +
5 DC-B	+
6 DC-C	++
7 DC-D	N.D.
8 DEA	++
N.D. (not determined)	

(4) Evaluation of antiserum (Western blotting)

5

10

15

20

35

40

45

50

55

[0172] For the confirmation of the binding of the antiserum prepared in the above (2) to Meltrins, Western blotting was carried out.

[0173] Mouse myoblast C2 was transformed by pBOSMel $\alpha\delta$ Pro(+) and pBOSMel β (+), which will be referred to hereinafter as "#9-3", and mouse myoblast C2 was transformed by pBOSMel $\alpha\delta$ MP(+), which will be referred to hereinafter as "#3-5."

[0174] The transformed C2 cells of 1 x 10⁷ cells were washed with PBS-(GIBCO BRL) and collected by centrifugation. The density of the collected cells was adjusted to 5 x 10⁶ cells/ml, mixed with a proteolysis inhibitor, C¢ mplete (Boehringer Manheim) in amount of one 25th of the volume of the cell-mixture, and mixed with SDS to a final concentation of 0.2%. After incubation at a room temperature for 30min, the cells were subjected to sonication at 4°C for 10sec (Isec x 10), and centrifuged. The resulting supernatant was collected and used as a cell lysate. Another cell lysate was prepared from fibroblast L929 (ATCC No.CCL-1) in the same way, and used as a negative control.

[0175] The resulting cell lysate (10μl) was mixed with an equiamount of a gel loading buffer (0.25M Tris-HCl, 2% SDS, 30% Glycerol, 0.01% BPB(pH 6.8)), the resulting solution (6μl) was applied to SDS-PAGE of 4-20T % (Tefco), and electrophoresed under 25mA at a room temperature for about one hour. After the completion of the electrophoresis, the contents were transferred to PVDF membrane (Millipore) under the conditions of 150mA, 4°C and 45min. The membrane was blocked by shaking in 4% skim milk (Meiji Milk Co.) at a room temperature for one hour, and each lane was cut. Each excised lane was soaked and shaked in antiserum (1ml) diluted 500 times with 50mM Tris-HCl (pH 7.2) comprising 0.05% Tween20 (referred to hereinafter as "T-TBS") and 4% skim milk at a room temperature for one hour. After the completion of the reaction, each lane was washed two times with T-PBS, soaked in 1ml of an anti-rat immunoglobulins antibody labelled with HRPO (Dako) diluted 500 times with T-PBS comprising 4% skim milk, and reacted at a room temperature for one hour. After washing five times with T-PBS, it was detected by ECL system (Amersham). The results are shown in Table 4. Bands were detected in the three kinds of the antiserums by the Western blotting.

TABLE 4 Reaction of antiserum with the cell lysate in Western blotting

Peptide antigens	Western blottting								
1 ProA	+								
2 MP-A	-								
3 MP-B	•								
4 DC-A	N.D.								
5 DC-B	N.D.								
6 DC-C	+								
7 DC-D	N.D.								

(continued)

Peptide antigens	Wes	stern blottting
8 DEA	+	
N.D. (not determined)		

(5) Preparation of monoclonal antibody

5

20

25

40

[0176] The antigens (ProA, MP-B, DC-C, DEA) (50µg each) were diluted with 400µl of physiological saline, and injected into the tail vein of the rats whose antibody titer had increased. Three days later, cell fusion was carried out by using myeloma P3X63Ag8U.1 according to the known method (Monoclonal antibody Jikken Sosa Nyumon (Guide of monoclonal antibody preparation), Tamie Ando and Jo Chiba, Koudan-sha Scientific). Six days later, the culture supernatant was collected and subjected to the plate assay according to the method of the above (3). The wells that showed reactivity with the peptide antigens were subjected to cloning by limiting dilution(Monoclonal antibody Jikken Sosa Nyumon (Guide of monoclonal antibody preparation), Tamie Ando and Jo Chiba, Koudan-sha Scientific). After cloning, the screening by the plate assay was performed again to give 27 clones of the hybridomas producing an anti-mouse Meltrin α monoclonal antibody which reacted with the peptide antigens. The results are shown in Table 5.

TABLE 5 Hybridomas producing anti-Meltrin peptides monoclonal antibody

Peptide antigens	Hybridoma No.	The number of Hb
ProA	F936	10
MP-B	F939	4
DC-C	F933	4
DEA	F934 .	

[0177] Purified antibodies were obtained from the thus established anti-Meltrin monoclonal antibody-producing hybridoma cell lines by the following method.

[0178] The hybridomas were cultured in RPMI1640 supplemented with 10% fetal bovine serum and 1ng/ml of human IL6 till a final density of 2 x 10⁵ cells/ml. The medium was then exchanged with a serum-free medium (Hybridoma-SFM, GIBCO BRL), and the culture was continued until the cells died. The resulting culture supernatant was filtered through filter paper for the removal of the cells, and subjected to purification by Protein G column (Prosep-G, Bioprocessing INC) as follows. The culture supernatant (1L) was applied to Prosep-G column (20ml) at a flow rate of 10ml/min, followed by washing with 0.1M phosphate buffer (pH 7.5) comprising 0.15M NaCl. After the absorbance at 280nm had decreased, the bound monoclonal antibody was eluted by 0.1M citric acid buffer (pH 3.0). After neutralization of the pH, the eluate was concentrated with DIAFLO (Grace Japan), and dialysed against 0.076M phosphate buffered saline (pH 6.4) comprising 0.45% NaCl. The concentration of the purified antibody was calculated on the basis of the absorbance at 280nm.

(6) Evaluation of monoconal antibody

[0179] The binding activity of 7 lots of the purified antibodies (10 µg/ml each) obtained in the above (5) to Meltrin was confirmed by Western blotting according to the method of the above (4) using the cell lysate of #9-3 cell. The results are shown in Fig. 19. The band of about 67kDa specific to the cell lysate of #9-3 cell was detected by the reaction with F933-4-3 (subclass IgG2a), F933-10-26 (subclass IgG2a), F934-17-6 (subclass IgG2a), F934-3-23 (subclass IgG2a), F934-4-33 (subclass IgG2a), F934-6-3 (subclass IgG2a), and F934-20-5 (subclass IgG2c). As these bands were not detected in the case of the cell lysate of L929 cell, it was confirmed that the monoclonal antibodies obtained in the above (5) were bound to Meltrin.

Example 11: Preparation of anti-mouse Meltrin monoclonal antibody

(1) Preparation of the antigen to be adminitered and immunization of rat

[0180] Rats were immunized with #9-3 and #3-5 cells as the antigen to be administered as follows. The cells used as the antigen to be administered were cultured in the absence of bFGF. First, the cells cultured in four dishes to a density of about 5 x 10⁵ cells /\phi 10cm dish were subcultured in 20 dishes to until the same density as the above, then again subcultured in 40 dishes (\phi 15cm) up to a density of about 5 - 6 x 10⁶ cells / dish, and futher cultured in a differentiation

medium (DMEM supplemented with 2% horse serum) for two days to finally form myotube. These cells were then scraped with a silicon rubber Policeman, washed two times with PBS, and suspended into the medium comprising 10% DMSO for storage at -80°C.

[0181] The #9-3 and #3-5 cells were suspended in physiological saline (200 μ l), mixed with an equiamount of Freund's complete adjuvant (DIFCO) and intraperitoneally administered into Wistar rat (5 weeks old, female) in an amount of 1 x 10⁷ cells/rat. The antigen was mixed with the same amount of Freund's incomplete adjuvant (DIFCO) and administered two weeks later in the same way as above.

(2) Evaluation of antiserum

10

30

35

40

50

[0182] After one week from the boosting, the blood was drawn from the eyeground of the rat, and a binding of antiserum to Meltrin was determined by using the cell extract according to the plate assay of Example 10 (3). The cell extracts of #9-3, #3-5 and L929 cells were prepared according to the method of Example 10 (4), except that NP-40 (Nacarai Tesque Co.) was used at a final concentration of 0.5% as a surfactant.

[0183] First, each cell extract was dilulted with PBS to a concentration of 40 μg/ml, each 50μl of which was separately poured into each well of an immuno plate (Maxisorp Nunc). After incubation at 56°C for 30min for binding of the antigen, the wells were washed five times with ion-exchanged water, blocked by 20 % Block Ace (Yukijirushi Milk Co.) / PBS 100μl, followed by incubation at a room temperature for 30min. After removal of the blocking agent, each antiserum (50μl) was added and incubated at 37°C for one hour. After two repeats of washing of the wells with the washing liquid, 50μl of an anti-rat immunoglobulins antibody labelled with peroxidase (Dako) and diluted 1,000 times with 10% Block Ace /PBS was added to each well followed by incubation at 37°C for one hour. Upon the completion of the reaction, the wells were washed five times with the washing liquid and two times with ion-exchanged water, and 50μl of 0.1M McIlvaine buffer (pH 5.0) comprising 3mg/ml of o-phenylene diamine and 0.027% hydro peroxide was added and reacted for 10 min. The reaction was terminated by the addition of 1N HCl (50 μl), and the absorbance at 490 nm was measured.

[0184] Western blotting was also carried out by using the cell extract of L4-3 described in the following (4) to confirm its binding to Meltrin. The results are shown in Table 6.

[0185] It was confirmed that the antiserum obtained from the rats immunized with #9-3 and #3-5 cells reacted with the corresponding cell extract, and were bound to Meltrin in the Western blotting.

TABLE 6 Reaction of antiserum of the rats immunized with #9-3 and #3-5 cells to Meltrin

L929	L4-3
-	+
-	+
	-

(3) Preparation of monoclonal antibody

[0186] The #9-3 and #3-5 cells (1×10^7 cells each) were suspended in physiological saline (200 μ I), and intraperitoneally administered into the rat whose antibody titer had increased. Three days later, cell fusion was carried out by using myeloma P3X63Ag8U.1 according to the known method (Monoclonal antibody Jikken Sosa Nyumon (Guide of monoclonal antibody preparation), Tamie Ando and Jo Chiba, Koudan-sha Scientific). Six days later, the culture supernatant was screened by its reactivity with the immobilized cell extracts. The wells that showed reactivity with the cell extracts were subjected to cloning by limiting dilution (Monoclonal antibody Jikken Sosa Nyumon (Guide of monoclonal antibody preparation), Tamie Ando and Jo Chiba, Koudan-sha Scientific). After cloning, the above screening was repeated to give 13 clones, 5 clone from the rat immunized with #9-3 (δ Pro; hybridoma No. F932) and 8 clones from the rat immunized with #3-5 (δ MP; hybridoma No. F937).

(4) Evaluation of monoconal antibody

[0187] The monoclonal antibodies F932-15-2 (subclass IgG1) and F937-9-2 (subclass IgG1) that showed a high reactivity with the cell extracts were evaluated.

[0188] First, the staining of myotube formed by C2 cells was examined by a cell immunofluorescence staining method. C2 cells were suspended in 10% FCS/DMEM at a density of 3 x 10⁴ cells/ml, each 100 µl of which was then separately poured into the wells of chamber slide (Lab-TEK, Nunc Co.). After the culture at 37°C and 5% CO₂ for two days, the medium was exchanged with 2% horse serum/DMEM. The cell staining was carried out by using myotube formed two

days later. The cells were washed two times with PBS⁻, and 4% formaldehyde was added followed by the reaction at a room temperature for 30min to fix the cells. The cells were washed three times with PBS⁻ and blocked with 20% Block Ace /T-PBS. After removal of the blocking agent, antibodies diluted to 10µg/ml with 20% Block Ace /T-PBS was added and reacted at a room temperature for one hour. After three repeats of washing of the wells with PBS⁻, an anti-rat immunoglobulins antibody FITC (Dako) diluted 20 times with 10% rabbit serum/T-PBS was added to each well followed by incubation a room temperature for one hour. After the completion of the incubation, the cells were washed three times with PBS⁻, and subjected to fluorescence microscopy. It was observed that myotube was stained by both the antibodies, but not stained by rat IgG (ZYMED) used as a negative control.

[0189] Next, L929 cells experssing mouse Meltrin α or β were prepared and subjected to cell staining for the purpose of confirmation of the specificity of the above antibodies. Thus, fibroblast L929 was transfected with the mixture comprising the plasmids pBOSMel α (+) and pBOSMel β (+) prepared in Example 4, and the plasmid pSV2NEO in a molar ratio of 12:12:1 by using LIPOFECTAMINE (Gibco BRL) according to its protocol to give L4-3 cells expressing mouse Meltrins α and β . Similarly, 1929 was transfected with the mixture comprising the plasmids pBOSMel β (+) and the plasmid pSV2NEO in a molar ratio of 20:1 to give L2-10 cells expressing mouse Meltrin β . Similarly, L929 was transfected with the plasmids pBOSMel α Pro(+) to give L8-5 cells expressing mouse Meltrin α Pro. The transfected cells were cultured in 10% FCS/DMEM and subcultured onto a chamber slide. The specificity of the antibodies was confirmed by cell staining using L929, L4-3, L2-10 and L8-5 cells. The results shown in Table 7 indicated that F932-15-2 was bound to Meltrins α and β , and F937-9-2 was bound to Meltrin α .

[0190] The hybridoma expressing the monoclonal antibody F932-15-2 was deposited with the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken 350 Japan) on October 3, 1996 under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and Regulation under accession numbers FERM BP-5687.

TΛ	DI.	7

Cell	Expression	F932-15-2	F937-9-2
L929	•	•	-
L4-3	α and β	+	+
L2-10	β	+	-
L8-5	α (δΡιο)	+	+

(5) Determination of neutralizing activity

25

30

35

[0191] The neutralizing activity of the monoclonal antibodies obtained in the above (3) was confirmed by their inhibition of the formation of myotube by C2 cells. C2 cells were cultured in a collagen-coated dish contianing 10% FCS/DMEM till 80% of confluence, followed by exchange of the medium with 2% horse serum/DMEM supplemented with 0 or $40\mu g/ml$ of the antibodies to be tested. The formation of byotube was then observed and the ratio of nuclei in the formed myotube was calculated. As seen from Fig.20, the formation of myotube on the day 2 was inhibited, showing that both F932-15-2 and F937-9-2 have the neutralizing activity.

Example 12: The activity of Meltrin neutralizing antibodies to inhibit the formation of bone resorption area (pit) in mouse unfractionated bone cells

[0192] Femur and tibia extracted from 13-day-old ICR mouse were crushed in MEM α medium (GIBCO) supplemented with 5% FBS. After being allowed to stand still for 2min, the precipitaed bone residues were removed. The supernatant of the suspending cells was adjusted to 1 x 10^7 cells/ml, 100μ l of which was then added to each well of a 96 well microplate provided with ivory fragments. The ivory fragments had been thinly sliced, punched into 6mm in diameter, washed with 70% ethanol and sterilized. The mouse Meltrin-neutralizing antibody (F932-15-2) obtained in Example 11, and rat IgG were diluted with MEM α medium (GIBCO) supplemented with 5 % FBS to final concentrations of 5, 50, and 500μ g/ml, 100μ l of which was then added to each well. After incubation at 37°C and 5% CO $_2$ for three days, the cells were removed with a scraper, and resorption area was stained with an acid hematoxylin solution (SIGMA) for about 7min and the number of the stained resorption area was caluculated using an ocular micrometer under a microscope by counting the number of squares wherein resorption fossa was contained.

[0193] The results are shown in Fig. 21, which demonstrates that the number of the formed resorption area was inhibited in a dose-depending manner by the mouse Meltrin-neutralizing antibody. Accordingly, it was suggested that the Meltrin-neutralizing antibody would affect directly or indirectly osteoclast and inhibit bone resorption.

Example 13: Serum Ca-decreasing activity of Meltrin-neutralizing antibody in mouse having enhanced bone resorption

[0194] Seven-week-old ICR mice (male) were fed for five days with low Ca feed with Ca content of 0.02% or less. The mouse Meltrin-neutralizing antibody (F932-15-2) obtained in Example 11 was injected into the tail vein of the mice (one group consisting of five mice) at doses of 0.1mg and 1mg per mouse). Rat IgG (Img per mouse) and phosphate buffer physiological saline were also injected as a control in the same way. Before injection and one day later, the blood was collected from the vein under eyes, and serum was separated. The value of Ca in the serum was then determined by an autoanalyzer (COBAS FARAII, ROCHE) using Ca determination kit (CalciumHR-II, WAKO Pure Pharmaceuticals). The results are shown in Fig.22.

10 [0195] As seen from Fig.22, the serum Ca value after one day from the injection in the groups treated with the mouse Meltrin-neutralizing antibody was lower than that of the groups treated with rat IgG or physiological saline. These results suggested that the Meltrin-neutralizing antibody would inhibit an unhealthly enhanced bone resorption due to hyperparathyroidism or malignant hypercalcemia.

5 Example 14: Preparation of cDNA fragment encoding human Meltrin α comprising transmembrane domain

[0196] A sense primer S-hMel α -TM5'was synthesized based on the partial cDNA sequence of human Meltrin α obtained in Example 8, and an antisense primer A-mMel α -3' was synthesized based on the cDNA sequence of mouse Meltrin α (see Table 1).

20 [0197] PCR was carried out by mixing the human placenta λgt11 cDNA library (Clonetech Co., code No. CLHL1008b) as a template, with S-hMelα-TM5' and A-mMelα-3' primers, Ex Taq polymerase (TaKaRa Co.,), and its reagents (TaKaRa Co.,), followed by 35 cycles of the reactions at 94°C for 30sec, 55°C for 30sec, and 72°C for one min. The base sequencing of the resulting amplified fragment (clone TM) suggested that the fragment was a human cDNA fragment corresponding to about 220 amino acids comprising the transmembrane domain of mouse Meltrin.

[0198] The obtained base sequence and its corresponding amino acid sequence are shown in Fig.23a - Fig.23b.

Example 15: Acute toxicity test

35

[0199] The mouse Meltrin-neutralizing antibody (F932-15-2) obtained in Example 11 was injected into seven-weekold ICR male mice (one group consisting of five mice) at doses of 1mg and 3mg permouse). Phosphate buffer physiological saline was also injected into a control group in the same way. Neither significant decrease of body weight nor side effect was observed in any group after the injection. No dead mouse was observed, either.

Reference Example 1: Preparation of monoclonal antibody recognizing human Meltrin

(1) Preparation of antibody using a peptide having the amino acid sequence derived from human Meltrin as an antigen

[0200] In consideration of the results obtained in Example 10, the sequence "GKVSKSSFAKCEMRDAKC" corresponding to DC-C in the amino acid sequence of human Meltrin α obtained in Example 8 was synthesized in the same way as in Example 10 (1), purified and conjugated with maleimidated KLH to give an antigen to be administered. 20μ g of the antigen was dissolved in 0.1ml of physiological saline and mixed with an equiamount of FCA followed by injection to ddy mouse (5 weeks old, female). The same amount of the antigen was mixed with FIA and injected two weeks later. The blood was collected from the eyeground one week later and antiserum was prepared. Evaluation of the reactivity of the resulting antiserum with the administered peptide according to the method of Example 10 (3) revealed its specific reactivity with the administered peptide. Accordingly, mouse, rat, hamster and the like are immunized with the peptide antigen, and monoclonal antibody may be prepared in the same manner as in Example 10 (5). Such antibody may also be used in Western blotting.

[0201] As it is estimated that the amino acid sequence in Fig.15a - Fig.15f is Meltrin α of a soluble type, an antibody, which may be effectively used in the determination of soluble Meltrin in the body, may be prepared by immunization of a peptide having the amino acid sequence adjacent to C-termial of the above sequence.

[0202] Similarly, antibodies recognizing human Meltrin β and Meltrin γ may be prepared by chemically synthesizing peptides having the amino acid sequences of suitable parts in the amino acid sequences in Fig.17a - Fig.17c or Fig.13a - Fig.13d and injecting the thus synthesized peptides into animals. In any case, the amino acid sequence will be selected from the extracellular domain.

[0203] For the preparation of an antibody specific to each one of Meltrins α, β and γ, the amino acid sequence should be selected from the parts with a low homology among them, and a peptide having the thus selected amino acid sequence is synthesized and injected to animals such as mouse, rat and hamster in the same way as in Example 10 (2).
[0204] In any case, monoclonal antibodies are prepared in the same way as in Example 10 (5).

(2) Preparation of anti-Meltrin monoclonal antibody using cells expressing human Meltrin as an antigen

[0205] DNA encoding the amino acid sequence wherein the amino acid sequence located downstream of the transmembrane domain shown in Fig.23a - Fig.23b is fused downstream of the sequence from the metalloproteinase or the disintegrin domain to the cysteine-rich region shown in Fig. 15a- Fig. 15f is prepared, and inserted into an expression vector pEFBOS, followed by transformation of C2 cells by the resulting vector. The transformant is treated as in Example 11 (1), and used as an antigen for immunization of animals such as mouse, rat and hamster. Antibodies recognizing human Meltrin α is screened as in Example 11 (2), and monoclonal antibodies are prepared as in Example 11 (3).

[0206] Similarly, DNA encoding the amino acid sequence shown in Fig.17a - Fig-17c or the sequence located downstream of the disintegrin domain of the above sequence is prepared, and inserted into an expression vector pEFBOS, followed by transformation of C2 cells by the resulting vector. The transformant is treated as in Example 11 (1), and used as an antigen for immunization of animals such as mouse, rat and hamster. Antibodies recognizing human Meltrin β is screened as in Example 11 (2), and monoclonal antibodies are prepared as in Example 11 (3).

[0207] Similarly, DNA encoding the amino acid sequence shown in Fig.13a - Fig.13d or the sequence located down-stream of the disintegrin domain of the above sequence is prepared, and inserted into an expression vector pEFBOS, followed by transformation of C2 cells by the resulting vector. The transformant is treated as in Example 11 (1), and used as an antigen for immunization of animals such as mouse, rat and hamster. Antibodies recognizing human Meltrin γ is screened as in Example 11 (2), and monoclonal antibodies are prepared as in Example 11 (3).

SEQUENCE LISTING

[0208]

20

25

30

35

40

45

50

- (1) GENERAL INFORMATION:
- (i) APPLICANT:
 - (A) NAME: MOCHIDA PHARMACEUTICAL CO., LTD.
 - (B) STREET: 7, Yotsuya 1-chome, Shinjuku-ku
 - (C) CITY: Tokyo
 - (E) COUNTRY: Japan
 - (F) POSTAL CODE (ZIP): 160
 - (ii) TITLE OF INVENTION: MELTRINS
 - (iii) NUMBER OF SEQUENCES: 28
 - (iv) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: EP 96935358.0

SEQ ID NO:1

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6915 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
- 55 (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA
 - (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO (vii) IMMEDIATE SOURCE

5

(B) CLONE: JM109(pBSMel α)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

	GCCI	AGAG:	TAG	CGCG	CGCGC	co co	ACGC/	ACACI	A CA	CGGG	GAGG	GGA	GAAA	GTT :	TTT.	TTGAA	60
10	AAA	ATGA	AAG	GCTA(BACTO	CG CT	rgc r	CAGC	3 AC	CCGG	GCGC	TGC	GCGA	GGG (GTC	CGGCA	120
	GAC	rcag(GGC .	AGTA	GAC!	CT C	ccca	AGCTY	C GG	CGCC	CGCG	TGG	GATG	CTG (CAGC	CTGGC	180
	CGCC	GGG(CCC	CCGA	AGCAC	C TO	GCAC	GCCA (GC(CGGC	GACA	ATG	GCA	GAG	CGC	CCG	235
												Met	Ala	Glu	Arg	Pro	
15	GCG	CGG	CGC	GCG	CCC	ccc	GCC	CGC	GCC	CTC	CTG	CTG	GCC	CTG	GCT	GGG	283
	Ala	Arg	Arg	Ala	Pro	Pro	Ala	Arg	Ala	Leu	Leu	Leu	Ala	Leu	Ala	Gly	
	GCC	CTG	CTG	GCG	CCC	CGT	GCA	GCC	CGA	GGG	ATG	AGT	TTG	TGG	GAC	CAG	331
20	Ala	Leu	Leu	Ala	Pro	Arg	Ala	Ala	Arg	Gly	Met	Ser	Leu	Trp	qaA	Gln	
20															~~~	000	
				TAC													379
	Arg	Gly	Ala	Tyr	Glu	Val	Ala	Arg	Ala	Ser	Leu	Leu	Ser	Lys	Asp	Pro	
	GGG	ATC	CCA	GGA	CAG	AGC	ATC	CCA	GCC	AAG	GAT	CAT	CCA	GAC	GTG	CTG	427
25	Gly	Ile	Pro	Gly	Gln	Ser	Ile	Pro	Ala	Lys	Asp	His	Pro	Asp	Val	Leu	

	ACT	GTG	CAA	CTG	CAG	CTG	GAG	AGC	CGA	GAC	CTG	ATC	CTC	AGC	CTG	GAA	475
	Thr	Val	Gln	Leu	Gln	Leu	Glu	Ser	Arg	qeA	Leu	Ile	Leu	Ser	Leu	Glu	
5	V CC	ስል ጥ	GAG	GGA	CTC	ATT	GCC	ААТ	GGC	TTC	ACG	GAG	ACC	CAT	TAT	CTG	523
5																Leu	323
	_			_													
					GAT												571
10	Gln	qaA	Gly	Thr	qaA	Val	ser	Leu	Thr	Arg	Asn	H18	Thr	Asp	нів	Cys	
10	TAC	TAC	CAT	GGA	CAT	GTG	CAA	GGA	GAT	GCT	GCA	TCA	GTG	GTC	AGC	CTC	619
																Leu	
15					GAT Asp											ACG	667
13	Ser	Int	Сув	Der	жыр	Leu	ALG	GIY	Deu	116	Mec	FIIG	GIU	ABII	Dyo	1111	
					CCA												715
	Tyr	Ser	Leu	Glu	Pro	Met	Lys	Asn	Thr	Thr	Asp	Ser	Tyr	Lys	Leu	Val	
20	CCA	CCT	GAG	AGC	ATG	ACG	AAC	ATC	CAA	GGG	CTG	TGT	GGG	TCA	CAG	CAT	763
	-				Met												
															0		
					CTC Leu											CAA Gln	811
25	ASII	гÀв	Ser	ASII	neu	1111	MEC	GIU	May	Val	361	FIO	GIY	1111	561	924	
																TAC	859
	Met	Arg	Ala	Arg	Arg	His	Lys	Arg	Glu	Thr	Leu	Lys	Met	Thr	Lys	Tyr	
	ር. Tr እ	GNG	CTG	CTT	ATT	CTC	CCA	GAC	AAC	AGA	GAG	John	CAG	AGG	CAA	GGA	907
30		_														Gly.	
					AAA											_	955
	rys	Азр	Leu	GIU	Lys	vai	гуя	GIII	Arg	rea	TTE	GIU	116	ALG	ABII	VIS	
35																GGA	1003
	Val	Asp	Lys	Phe	Tyr	Arg	Pro	Leu	Asn	Ile	Arg	Ile	Val	Leu	Val	Gly	
	GTY:	CAA	GTG	TGG	AAT	GAC	ልጥሮ	GAC	444	TGC	тст	АТА	AGC	CAG	GAC	CCA	1051
					naA												1031
40				_													/225
					CAT His												1099
	Pne	THE	Arg	Den	nis	GIU	Pne	rea	жър	IIÞ	Arg	Lys	116	Lys	LEU	Deu	
					CAC												1147
45	Pro	Arg	Lys	Ser	His	Asp	Asn	Ala	Gln	Leu	Ile	Ser	Gly	Val	Tyr	Phe	
	CAA	GGA	ACC	ACC	ATC	GGC	ATG	GCA	CCC	ATC	ATG	AGC	ATG	TGC	ACT	GCA	1195
					Ile												
												a				COT	
50					GGA Gly												1243
	GIU	GIU	SEL	GIA	GIY	AGT	AGT	riel	voh	1140	JE4	, L	562			~~;	
					TTG												1291
	Ala	Ala	Val	Thr	Leu	Ala	His	Glu	Leu	Gly	His	Asn	Phe	Gly	Met	Asn	
55	САТ	GAC	ארש	CTY	GAG	ACC	GGC	ፕሮር	AGC	TGC	AGA	ATG	GCC	GCA	GAG	AAA	1339
		~~~	3												-	-	

	His	Asp	Thr	Leu	Glu	Arg	Gly	Сув	Ser	Суз	Arg	Met	Ala	Ala	Glu	Lys	
	GGA	ccc	TGC	ATC	ATG	AAC	CCG	TCC	ACG	GGG	TTC	CCA	TTC	CCC	ATG	GTG	1387
5																Val	
																GGC Gly	1435
	Pue	ser	Ser	cys	Ser	Aty	Був	vab	Deu	GIU	MIG	361	Tien	GIU	Lys	GLY	
10	ATG	GGG	ATG	TGC	CTC	TTC	AAC	CTA	CCA	GAG	GTC	AAG	CAG	GCC	TTT	GGG	1483
	Met	Gly	Met	Сув	Leu	Phe	Asn	Leu	Pro	Glu	Va1	Lys	Gln	Ala	Phe	Gly	
	-00			mam	GGA		000	m> m	ama	<b>~~~</b>	CNC	CCN	GD D	C3.C		CNC	1531
																Asp	1531
15	<b>U</b> -y	9	2,5	O, O	07		<b>4-7</b>	-1-				3			-2-		
																ACC	1579
	Сув	Gly	Glu	Pro	Glu	Glu	Сув	Thr	Asn	Arg	Сув	Cys	Asn	Ala	Thr	Thr	
	aven.	እርጥ	C-IV2	A A G	CCA	GAT	CCT	CTYC	ተርር	cca	CAC	GGG	CAG	TGC	TGT	GAA	1627
20																Glu	202,
	•																
																AAC	1675
	Asp	Сув	Gln	Leu	Lys	Pro	Pro	GIA	Thr	Ala	Cys	Arg	GIA	ser	ser	Asn	
25	TCC	TGT	GAC	CTC	CCA	GAA	TTC	TGC	ACA	GGG	ACT	GCC	CCT	CAC	TGT	CCA	1723
					Pro												
																GGT Gly	1771
30	ATA	ABII	val	Tyr	rea	птв	мър	GIY	HID	PIO	Cys	GIII	GLY	AGT	rop	GLY	
																ACG	1819
	Tyr	Сув	Tyr	neA	Gly	Ile	Cys	Gln	Thr	His	Glu	Gln	Gln	Cys	Val	Thr	
	CTC	TCC	CGN	CCA	GGT	CCT	222	CCG	CCT	ርርም	GGC	יאדמ	TGC	باحلحل	GAG	CGA	1867
35																Arg	200,
		-															
					GGA												1915
	Val	Asn	Ser	Ala	Gly	Asp	Pro	TYY	GIA	Asn	Сув	GIA	rys	Asp	ser	гув	
40	AGC	GCC	TTC	GCC	AAA	TGT	GAG	CTG	AGA	GAT	GCC	AAG	TGT	GGG	AAA	ATC	1963
					Lys												
										cma			100			Carro	2011
					GGT Gly												2011
45	GIII	Cys	9111	GLY	GIY	774		<b>—</b> э		742		<b>U</b> _,					
					AAT												2059
	Ser	Ile	Glu	Thr	asa	Ile	Pro	Gln	Gln	Glu	Gly	Gly	Arg	Ile	Leu	Сув	
	CCC	ccc	እሮር	_C አጥ	GTG	TAC	عملمد	GGT	GAT	GAC	DTG.	CCA	GAC	CCA	GGG	CTT	2107
50					Val												
		_															
					ACA												2155
	val	Leu	ATA	GīÅ	Thr	гÀа	Сув	ATS	GIU	GIĀ	гåа	TTE	CAR	Leu	MBII	щā	
55	CGA	TGT	CAG	AAT	ATC	AGT	GTC	TTC	GGC	GTT	CAC	AAG	TGT	GCC	ATG	CAG	2203
					Ile												

	TGC	CAC	GGC	CGA	GGG	GTA	TGT	AAC	AAC	AGG	AAG	AAT	TGC	CAC	TGT	GAA	2251
	Сув	His	Gly	Arg	Gly	Val	Сув	Asn	Asn	Arg	Lys	Asn	Сув	His	Cys	Glu	
5																AGC	2299
	Ala	His	Trp	Ala	Pro	Pro	Pne	Сув	Asp	rys	Phe	GIÅ	Phe	GLY	GIA	Ser	
	N.C.N	GNC	ACT	CCT	ccc	ATC	» CC	CDD	CCA	CAT	220	CNG	ccc	TTY	እርጥ	CTA	2347
						Ile											2347
10		.wp		<b>-</b>			3	<b>U</b>		- LUP		<b>V</b> 2	0_,				
	GGA	ATC	CTG	GTG	AGC	ATC	CTG	TGT	CTG	CTT	GCT	GCT	GGA	TTT	GTG	GTG	2395
	Gly	Ile	Leu	Val	Ser	Ile	Leu	Сув	Leu	Leu	Ala	Ala	Gly	Phe	Val	Val	
15						ACG											2443
15	Tyr	Leu	Lys	Arg	Lys	Thr	Leu	Met	Arg	Leu	Leu	Phe	Thr	His	Lys	Lys	
	. 200	300	N TOC	C22	270	CTA	NCC.	TOT	CTC	CNC	ССТ	TOC	cca	202	ccc	N.CTT	2491
						Leu											2431
		***		024	2,5			<b>C</b> 7.0	142	0			,			JC2	
20	GGC	CCT	CAC	CTT	GGC	CAG	GCT	CAC	CAC	ACC	CCC	GGG	AAA	GGC	CTG	CTG	2539
	Gly	Pro	His	Leu	Gly	Gln	Ala	His	His	Thr	Pro	Gly	Lys	Gly	Leu	Leu	
						CAT											2587
25	Met	Asn	Arg	Ala	PIO	His	Pne	ASN	Thr	Pro	гÀв	двр	Arg	HIB	ser	rea	
	AAA	TGC	CAG	AAC	ATG	GAC	ATC	AGC	AGG	CCC	CTC	GAC	GCT	CGA	GCC	GTC	2635
						Asp											
	•	•				-						-					
						CCT											2683
30	Pro	Gln	Leu	Gln	Ser	Pro	Gln	Arg	Val	Leu	Leu	Pro	Leu	His	Gln	Thr	
				000		~~~	~~~			000		000	~~~	1 CM		CC2	2721
						GGC Gly											2731
	710	AL 9	<i>_</i>		501	O ₂			~-3		204		744			~~u	
35	GTC	AGG	CAG	GCC	CAG	GGC	ATT	CGA	AAA	CCC	AGT	CCT	CCT	CAG	AAG	CCT	2779
	Val	Arg	Gln	Ala	Gln	Gly	Ile	Arg	Lys	Pro	Ser	Pro	Pro	Gln	Lys	Pro	
							_						_				
						CTG											2827
40	Leu	Pro	ALA	Asp	Pro	Leu	ser	Arg	Thr	ser	Arg	Leu	Thr	ser	ATS	Leu	
	GTG	AGG	ACC	CCA	GGG	CAG	CAG	GAA	ССТ	GGG	CAC	ന്ദ്രവ	CCA	GCC	ccc	ATC	2875
						Gln											
					_					_							
																TAT	2923
45	Arg	Pro	Ala	Pro	Lys	His	Gln	Val	Pro	Arg	Pro	Ser	His	Asn	Ala	Tyr	
			masa				C 2 C C	an					~~~	5 5 CT	AIAIN- (	73.0	2070
			TGAC	iaagc	CA G	CCCA	GACC	.6 61	CCTC	AACA	GIU	iaaga	CAG	AAGI	1160	LAC	2979
	Ile	4y3															
50	TATO	TTC	IGC 1	CCAT	TGGA	G TI	GTTG	TTGI	ACC	AACT	TTC	CGAG	TTTC	TA A	AGTO	TTTAA	3039
	AACA	CCAT	TC 1	CTCC	AGAC	C C1	GGAG	CCAC	TGC	CATC	:GGT	GCTG	TGCI	GT G	GTGC	TTTGT	3099
																CAGGG	3159
					_	-	-						-			CGTGC	3219
55																IGTTTA CCTAC	3279
	CITI	CTAI	TTC A	MGGC	CITA	AT CC	<b>GAAA</b>	ATAG	CIC	ייים.	CCT	TCCC	AAGG	iCI, G	LLAI	rggtac	3339

	CAGACACACA	GCTCAGGACA	CCCCAGGGAG	AACCTGGCAT	GGGTTTTCTT	TGTTTGCTTT	3399
	CATTTTATCT	TTTATATTTT	GGTATCCCTA	TCTTGGGTTG	TAGCCAGGGC	CTTCAGGAAG	3459
	GTCTTGGGCC	ACTGCATGCT	AATGGCCTTC	AGGTCCTGCA	CCCTGAAGCT	CTCAGACAAC	3519
5	AAGTAGGATC	TGCTTTCTAG	CCAGCAGCTT	TGGAGAGAAC	CTGGGGTACT	GAAAAGAAGG	3579
	TTTGGGGTGT	<b>GGTTATACCA</b>	GGATGGAGAC	TGGAATCCTA	ATCTGGGCAA	ACATCTGACC	3639
	TTGAGCTGAG	CAGCCATGAG	CACCTCTAGG	AAGCAAGGAC	GGCTGAGGTG	CTGCACAAGG	3699
	CTCTGCTTTG	AGAGCTGGCA	GGGGCTTCTC	TCTGGCTGCC	CTTTGCAGAG	TGCTAGCTGG	3759
	CATGGCATGT	TGTTTACATC	GGGAACAGTG	GTGTTTCTAC	AAGAAAGCCA	CTGCCTGGGC	3819
10	ACTGCAGACC	TCCGTCTCCT	GCCCATTTAG	AGCTAAGCAA	ATTACCACAT	TGTCTTCTGG	3879
		AATGACCCTG					3939
		TGTGAACTAG				-	3999
		CAAGGCTCGA					4059
		-				TTATTAGTAG	4119
15		AAGCACCCAA					4179
	• • • • • • • • • • • • • • • • • • • •	GACTCGGTCT					4239
	•••	GTTTCCGAAT		-			4299
						CAAATTITCA	4359
		AACCCCATNG					4419
20						ACGACAGAAG	4479
		GCAGTCCTCT					4539
	•				-	CAGATAAGGA	4599
		CTTTCTAGAA					4659
		TAAATCCACG					4719
25		CAGAGACAAT					4779
		AGTAGTTGTA					4839
		TGTCTTGTGA					4899
	********	TGATGGGGAG					4959
		GAGAGAGTGT					5019
30	-	CTAAGGATGC					5079
55		CATGCATTCA					5139
		GTGTTACTAA					5199
		AACTTCAATT					5259
		TAGTTTCTCT					5319.
35		ATTTGCAGTC					5379
33		GGGAGACCAT					5439
		CGTCCTCCTT					5499
		GCAAGTGGGA					5559
		GTCCTGAGTC					5619
40		TGCATGCCGA					5679
40		CAGAAACCAC					5739
		GAGGAGACAT					5799
		CATTGTCCCC					5859
						GGTGCTANAA	5919
45						AGCAGATGTT	5979
45	TACTGAGCAC	TCTGAGCCAG	AAGCACCCCG	ACAACCAGGA	GGACGATNGC	TGGGCAGTAG	6039
						ATTCCCTTCT	6099
						CAGACGGCCT	6159
						TCCCCAAGAA	6219
50	CTGGTTTTTA	AACACTATGA	CAAGTAGAAG	AGGGTGTCAC	AGAGGCCATT	TTTTTTCTTT	6279
50						TTTACACAAC	6339
						TIGCCTICCT	6399
						CTGTCATCCA	6459
						GCTAACCTTC	6519
						AGCAACCTCC	6579
55						ACAGTGACCT	6639

699
759
819
879
915

# 10 SEQ ID NO:2:

5

15

20

25

30

35

40

45

50

55

# (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 903 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

31 .

Met Ala Glu Arg Pro Ala Arg Arg Ala Pro Pro Ala Arg Ala Leu Leu Leu Ala Leu Ala Gly Ala Leu Leu Ala Pro Arg Ala Ala Arg Gly Met Ser Leu Trp Asp Gln Arg Gly Ala Tyr Glu Val Ala Arg Ala Ser Leu Leu Ser Lys Asp Pro Gly Ile Pro Gly Gln Ser Ile Pro Ala Lys Asp His Pro Asp Val Leu Thr Val Gln Leu Gln Leu Glu Ser Arg Asp Leu Ile Leu Ser Leu Glu Arg Asn Glu Gly Leu Ile Ala Asn Gly Phe Thr Glu Thr His Tyr Leu Gln Asp Gly Thr Asp Val Ser Leu Thr Arg Asn His Thr Asp His Cys Tyr Tyr His Gly His Val Gln Gly Asp Ala Ala Ser Val Val Ser Leu Ser Thr Cys Ser Asp Leu Arg Gly Leu Ile Met Phe Glu Asn Lys Thr Tyr Ser Leu Glu Pro Met Lys Asn Thr Thr Asp Ser Tyr Lys Leu Val Pro Ala Glu Ser Met Thr Asn Ile Gln Gly Leu Cys Gly Ser Gln His Asn Lys Ser Asn Leu Thr Met Glu Asp Val Ser Pro Gly Thr Ser Gln Met Arg Ala Arg Arg His Lys Arg Glu Thr Leu Lys Met Thr Lys Tyr Val Glu Leu Val Ile Val Ala Asp Asn Arg Glu Phe Gln Arg Gln Gly Lys Asp Leu Glu Lys Val Lys Gln Arg Leu Ile Glu Ile Ala Asn His Val Asp Lys Phe Tyr Arg Pro Leu Asn Ile Arg Ile Val Leu Val Gly Val Glu Val Trp Asn Asp Ile Asp Lys Cys Ser Ile Ser Gln Asp Pro Phe Thr Arg Leu His Glu Phe Leu Asp Trp Arg Lys Ile Lys Leu Leu Pro Arg Lys Ser His Asp Asn Ala Gln Leu Ile

40

5

10

15

20

25

*3*0

*3*5

45

50

55

5

10

15

20

25

30

*3*5

40

45

50

55

Ser Gly Val Tyr Phe Gln Gly Thr Thr Ile Gly Met Ala Pro Ile Met Ser Met Cys Thr Ala Glu Gln Ser Gly Gly Val Val Met Asp His Ser Asp Ser Pro Leu Gly Ala Ala Val Thr Leu Ala His Glu Leu Gly His Asn Phe Gly Met Asn His Asp Thr Leu Glu Arg Gly Cys Ser Cys Arg Met Ala Ala Glu Lys Gly Gly Cys Ile Met Asn Pro Ser Thr Gly Phe Pro Phe Pro Met Val Phe Ser Ser Cys Ser Arg Lys Asp Leu Glu Ala Ser Leu Glu Lys Gly Met Gly Met Cys Leu Phe Asn Leu Pro Glu Val Lys Gln Ala Phe Gly Gly Arg Lys Cys Gly Asn Gly Tyr Val Glu Glu Gly Glu Glu Cys Asp Cys Gly Glu Pro Glu Glu Cys Thr Asn Arg Cys Cys Asn Ala Thr Thr Cys Thr Leu Lys Pro Asp Ala Val Cys Ala His Gly Gln Cys Cys Glu Asp Cys Gln Leu Lys Pro Pro Gly Thr Ala Cys Arg Gly Ser Ser Asn Ser Cys Asp Leu Pro Glu Phe Cys Thr Gly Thr Ala Pro His Cys Pro Ala Asn Val Tyr Leu His Asp Gly His Pro Cys Gln Gly Val Asp Gly Tyr Cys Tyr Asn Gly Ile Cys Gln Thr His Glu Gln Gln Cys Val Thr Leu Trp Gly Pro Gly Ala Lys Pro Ala Pro Gly Ile Cys Phe Glu Arg Val Asn Ser Ala Gly Asp Pro Tyr Gly Asn Cys Gly Lys Asp Ser Lys Ser Ala Phe Ala Lys Cys Glu Leu Arg Asp Ala Lys Cys Gly Lys Ile Gln Cys Gln Gly Gly Ala Ser Arg Pro Val Ile Gly Thr Asn Ala Val Ser Ile Glu Thr Asn Ile Pro Gln Glu Gly Gly Arq Ile Leu Cys Arg Gly Thr His Val Tyr Leu Gly Asp Asp Met Pro Asp Pro Gly Leu Val Leu Ala Gly Thr Lys Cys Ala Glu Gly Lys Ile Cys Leu Asn Arg Arg Cys Gln Asn Ile Ser Val Phe Gly Val His Lys Cys Ala Met Gln Cys His Gly Arg Gly Val Cys Asn Asn Arg Lys Asn Cys His Cys Glu Ala His Trp Ala Pro Pro Phe Cys Asp Lys Phe Gly Phe Gly Gly Ser Thr Asp Ser Gly Pro Ile Arg Gln Ala Asp Asn Gln Gly Leu Thr Val Gly Ile Leu Val Ser Ile Leu Cys Leu Leu Ala Ala Gly Phe Val Val Tyr Leu Lys Arg Lys Thr Leu Met Arg Leu Leu Phe Thr His Lys Lys Thr Thr Met Glu Lys Leu Arg Cys Val His Pro

Ser Arg Thr Pro Ser Gly Pro His Leu Gly Gln Ala His His Thr Pro Gly Lys Gly Leu Leu Met Asn Arg Ala Pro His Phe Asn Thr Pro Lys Asp Arg His Ser Leu Lys Cys Gln Asn Met Asp Ile Ser Arg Pro Leu Asp Ala Arg Ala Val Pro Gln Leu Gln Ser Pro Gln Arg Val Leu Leu Pro Leu His Gln Thr Pro Arg Ala Pro Ser Gly Pro Ala Arg Pro Leu Pro Ala Ser Pro Ala Val Arg Gln Ala Gln Gly Ile Arg Lys Pro Ser Pro Pro Pro Gln Lys Pro Leu Pro Ala Asp Pro Leu Ser Arg Thr Ser Arg Leu Thr Ser Ala Leu Val Arg Thr Pro Gly Gln Gln Glu Pro Gly His Arg Pro Ala Pro Ile Arg Pro Ala Pro Lys His Gln Val Pro Arg Pro Ser His Asn Ala Tyr Ile Lys 698

SEQ ID NO:3:

25

5

10

15

20

30

35

45

50

55

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 6345 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vii) IMMEDIATE SOURCE

(B)CLONE: JM109(pBSMel β)

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

	GGC	CGGG	GC 2	AGGC	aatg	GC A	GGGG	ATGT	G TG	ATTG	CGGA	CAG	TGAG	AGG (	SCCG:	ltgct.	A	60
5		ATG (															:	107
		CTG Leu					-										;	155
10		AGA Arg														ATA Ile	:	203
15		CCT Pro														CCA Pro	:	251
		AGA Arg														CTA Leu	:	299
20		CTG Leu															:	347
25																		

5

	TGC	TAC	ACT	GCA	agt	GGC	AAT	CCT	CAA	ACC	AGC	ACG	CTG	AAG	TCT	GAG	395
	Сув	Tyr	Thr	Ala	Ser	Gly	Asn	Pro	Gln	Thr	Ser	Thr	Leu	Lys	Ser	Glu	
5	GAT	CAC	TGC	TTT	TAC	CAC	GGG	ACT	GTG	AGG	GAC	GTG	GAT	GAG	TCC	AGT	443
																Ser	
	GTC	ACG	CTC	AGC	ACC	TGC	CGG	GGA	ATT	AGA	GGA	CTG	ATT	ATA	GTG	AGA	491
10					-	_			_							Arg	
	AGT	, DDC	CTC	AGC	TAC	ATC	ATC	GAG	ccc	GTC	CCT	AAC	AGC	GAC	AGC	CAA	539
																Gln	
	CAC	ርርጥ	ייייי ע	тас	DCD	TCC	GAA	СУТ	CTC	ACG	CALC	CCC	CCG	GGG	AAC	TGT	587
15																Сув	50,
	occ	محمد	CNG	CAC	TCC	aca	ccc	»cc	TCG	AAG	GAC	TCC	ccc	ملحلت	CAG	TTT	635
																Phe	033
20	202	O N TO	CNC	NCC.	222	220	CAA	CCT	ccc	n C'h	בעדת	444	cca	CAA	CAT	CTA	683
																Leu	003
	a.a	mam	<b>.</b>	220	T. C	CTC	ana.	CTV	ma c	~~~	CTC.	CCT	ር እጥ	ጥእጥ	CCA	GAG	731
oe.																Glu	,31
25	~~~	<b>a.</b> a	330	225	CCA	CAT	CNG	CAC	Cam	CCC	N.C.C	222	CCC	A A C	CTC	A TYC	779
						His											,,,
			-		<b></b>	amm	- C - T			<b></b>	000	mcc.	CMC	220	3 m/c	CCA	827
30						Val										CGA Arg	827
																	075
		-				Leu										GAA Glu	875
35																	
						TAC Tyr											923
						CAG Gln											971
40	_	-															1010
						GGC Gly											1019
	_	_															1067
45						CAG Gln											1067
		_			•			_									1116
						GCC Ala											1115
50																	1167
50						GAT Asp											1163
						_											
						ATG Met											1211
	•			-							_					-	

	GTG	TTC	AGT	TGG	TGT	AAC	AGG	AAG	GAG	CTG	GAC	AGG	TAT	CTG	CAG	ACA	1259
	Val	Phe	Ser	Trp	Сув	Asn	Arg	Lys	Glu	Leu	qeA	Arg	Tyr	Leu	Gln	Thr	
5	GGA	GGA	GGG	ATG	TGT	CTC	TCC	AAC	ATG	CCG	GAC	ACT	AGG	ACG	CTG	TAT	1307
	Gly	Gly	Gly	Met	Сув	Leu	Ser	Asn	Met	Pro	Asp	Thr	Arg	Thr	Leu	Tyr	
	GGA	GGC	CGG	AGG	TGT	GGC	AAC	GGG	TAC	CTG	GAA	GAC	GGT	GAA	GAA	TGT	1355
10	Gly	Gly	Arg	Arg	Сув	Gly	Asn	Gly	Tyr	Leu	Glu	qaA	Gly	Glu	Glu	Сув	
	GAC	TGT	GGA	GAA	GAG	GAG	GAA	TGT	AAG	AAC	CCT	TGC	TGC	AAT	GCC	TCC	1403
	qaA	Сув	Gly	Glu	Glu	Glu	Glu	Сув	Lys	Asn	Pro	Сув	Сув	Asn	Ala	Ser	
	AAC	TGC	ACT	CTG	AAG	GAA	GGG	GCA	GAG	TGT	GCC	CAT	GGT	TCC	TGC	TGC	1451
15 .	Asn	Сув	Thr	Leu	Lys	Glu	Gly	Ala	Glu	Сув	Ala	His	Gly	Ser	Сув	Сув	
	CAC	CAG	TGC	AAG	CTG	GTG	GCT	CCT	GGA	ACC	CAG	TGT	CGG	GAG	CAG	GTT	1499
			_	-		Val			_								
20						CCC											1547
			_			Pro											
																CAG	1595
25						Gln											
						GGC											1643
		_				Gly											1601
						GGA											1691
30						Gly											1739
	-	-				Gly											1133
05	_					AAG			-	_		_		_			1787
35						Lys											_,,,
		_		_		ACC											1835
						Thr											
40			_														
	TCT	ATT	GAC	ACC	ACC	ATC	ACC	TTG	AAC	GGG	AGG	CGG	ATC	CAC	TGT	CGG	1883
			_			Ile					_	_					
						CGG											1931
45						Arg											
						CTG											1979
			_		•	Leu						•	•	-			200
50						GGG											2027
			_			Gly		_									20-5
						AAG											2075
55		_				Lys											24.5
	AAG	AAC	TGT	CAT	TGC	TTC	CCL	GGC	TGG	TCT	CCA	CCT	TIC	TGT	AAC	ACC	2123

	Lys	Asn	Сув	His	Сув	Phe	Pro	Gly	Trp	Ser	Pro	Pro	Phe	Сув	Asn	Thr	
	cce	GGA	GAT	GGT	GGC	AGC	GTC	GAC	AGT	GGT	CCT	TTG	CCC	CCT	AAG	AGT	2171
5					Gly												
		_	_	Ī	_					_					-		
																GCA	2219
	Val	Gly	Pro	Val	Ile	Ala	Gly	Val	Phe	Ser	Ala	Leu	Phe	Val	Leu	Ala	
10	CTT	CTC	GTG	CTA	CTG	ጥርብ	CAC	TGC	TAC	AGA	CAG	AGC	CAC	222	CTG	ccc	2267
10					Leu												
	•					-		•	•	•				•		•	
					CTC												2315
	Lys	Pro	Ser	Ala	Leu	Pro	Phe	Lys	Leu	Arg	His	Gln	Phe	Ser	Сув	Pro	
15	سال	AGG	GTA	тст	CAG	AGT	CGT	GGA	ACT	GGC	CAT	GCC	AAC	CCA	ACT	TTC	2363
					Gln												
		•					-	-		•							
					CCC											_	2411
20	Lys	Leu	Gln	Thr	Pro	Gln	Gly	Lys	Arg	Lys	Val	Thr	Asn	Thr	Pro	Glu	
	TCC	כידר	CGG	DAG	CCC	TCC	CAC	CCC	ССТ	CTC	CGG	ccc	CCT	CCA	GAC	TAC	2459
					Pro												
			_	•													
25					TCG												2507
	Leu	Arg	Val	Glu	Ser	Pro	Pro	Ala	Pro	Leu	ser	Ala	H18	Leu	Asn	Arg	
	GCT	GCT	GGG	AGC	TCC	CCA	GAA	GCT	GGG	GCT	CGA	ATA	GAA	AGA	AAG	GAG	2555
					Ser												
30			_														
					CCT												2603
	Ser	Ala	Arg	Arg	Pro	Pro	Pro	ser	Arg	PIO	met	PIO	PIO	ALA	PLO	ASI	
	TGC	CTA	CTG	TCC	CAG	GAC	TTC	TCC	AGG	CCT	CGA	CCA	CCT	CAG	AAG	GCA	2651
35																Ala	
					CCG Pro												2699
	Leu	PLO	ALA	ABII	PIO	Val	PLO	GLY	<b>G</b> 211	AL 9	****	GLY	PIU	~4	Jer	GLY	
40	GGC	ACC	TCC	CTG	CTT	CAG	CCC	CCT	ACT	TCT	GGT	CCT	CAG	CCC	CCC	AGG	2747
40	Gly	Thr	Ser	Leu	Leu	Gln	Pro	Pro	Thr	Ser	Gly	Pro	Gln	Pro	Pro	Arg	
	000	003	CON	CTC	CCT	ملحلت	CCA	226	ста	ccc	GAG	TAC	CGA	TCA	CAG	ngg	2795
					Pro												2755
			7124					-3-				-1-	3				
45					ATT					TAGA	LAGTO	TC G	AGAA	GTT	C		2842
	Val	Gly	Ala	Ile	Ile	Ser	Ser	Lys	Ile								
	Trefference of the Control	MIN CYC	አጥ ር	CAAC	! እርግፕር	יר פס	יארננ	יראיני	CAL	CGTC	מרצי	AAGI	מאמ	יכפ נ	יושויי	TCACC	2902
																CTTAC	2962
50																CAAGTG	3022
																GGGAA	3082
																GAGCT	3142 3202
																GGAAC TACTG	3262
55																CATGT	3322

	AATAAGCCAT	GCTCCCCTCC	CCTGCCTTTC	TTCACATTCC	CACTCCCATA	TTTACACGGG	3382
	TCACTCTGAC	TCAGACAGGT	ACTATTTGTA	AGTAGCATAG	ACAGCAGGGG	GGTGGGGTGG	3442
	TCAACCTGTG	TCCCCTCTGA	GCCGTTATGC	CAAAGGTCAC	TAAGGACATT	TAGAATCCCC	3502
5	ATCCATCCAT	CCATCCATCC	ATCCATCCAT	CCATTCATCC	ATCCCCAGTG	TTCCATGTGT	3562
	CACCTTCTCC	TTTTCCAGCA	TCCCTATCCT	ATGGTGCTTT	GGTGGTGAAC	TATGGCAGTC	3622
	CTGACTTGCT	GATGACCATA	TGCTGGTGAC	CTACAAATCG	GGATCCTGCC	ATATGGGGTC	3682
	GCCACTGGAC	TTTCTGCACT	GGTTCTCAAG	AGCGTTGAGC	CGAGTGGGCG	TGTATGTTTG	3742
	TGTGTGTGTG	TGTGTGTGTG	TGTGTGTGTG	TGTGTGTGTG	TGTGTGTGTG	TGTGTGTGTG	3802
10	AAAGAGACAG	AGGCAATGAG	AGAGACAGAC	<b>ATGCAGGCAG</b>	GCCGACAGCT	CTGCATGTAC	3862
	TIGIGITITA	CGGCCTCAAG	CAGTATAAGG	GACCTCCTCC	TTATTTCTGA	CTCATATCTA	3922
	AGTAAGGTTC	CCCAGGACAG	CCACAGCTGT	ACTGAGGGGG	GCTGACATGT	TTGGCATCCT	3982
	GGCTATAGTA	TTGTATACAC	AGGGCCACCA	GCCCCGCCCT	AGTGGTCAGC	TCTGAGGGGG	4042
	GACTGGTGAC	TCTGAACAGA	TCGATGTCAA	CAGCCATGGT	GAACCAGATC	TGGGCAGGGT	4102
15	TCCCCAAACT	CTATTCAACC	AGAGTTTTAT	CACGCACTCA	TCGGGTCTCT	CCTGGTTGCT	4162
	GCCCCGAGGT	GATCGTCATG	GAAAATGCTG	AGAAGGTGGG	AATGGGATGG	GGTGGACCTT	4222
	CTCTTGCTTG	GTGCTCCGCT	ATTTGGAACA	GTTCTTACAC	ATTTGCTGGG	CCTGGCCTCT	4282
	GAGAGGCCAT	CTTCCACCCC	CAGAAAGGTG	CTAATGGCAC	TGCAGAGGGC	TCTCTAGGGG	4342
	CCTCCCCGCC	CCAACAGCAA	GCAGTTGTTA	GCTCTTGGAA	CCCTCCAGAG	GAAGAGGCAA	4402
20	GCGTTTGACT	TCCCCTTTAC	CACCTGAGGC	CTCCTTATAT	CTCTTCCCAG	AGTAAGCTTT	4462
	GGGATTGTAG	ACATGTGGGA	GCTATGACAG	ACGTGGCCTG	GGGTAGAAAG	ATCTCAGGAA	4522
			GGTGACCGTG				4582
	GTCCTATATC	AGTTTCTCTT	TTGTGTGCTT	TACCAAGTGG	CCGGTGACTA	CAGGCCACCC	4642
	CGATTCTCAC	CACAAAGTTA	GAAACCCTCC	ACTTTCTGTC	CCTTGAACCA	TATCAGAAAA	4702
25	AGACCCATTT	CCTTGCTCTT	TGGTAATCAC	TTCTGTTTTT	TCTTCTTCAT	TACTGTGCTA	4762
	CCACCTCCAT	CCCATGACAT	TATTCTGTGA	GTGTAAGAGG	ACGGTGTTTT	TTATCTTGGG	4822
	AGAATGTCGG	CAGCTGCTCT	ACACACAACT	TCACTCAAGG	CTTTGTCTCC	AGAGGCCAGC	4882
	TAGGCTGTCA	CAGGCAGGAA	TCCCTTCCCA	TCTGCTTTGT	GAAGGGTCCC	ATACAGGTGT	4942
	ATCTAGACTT	CAAGGACAGG	GTTTGTCTCA	CAGGATTGTC	<b>ACTTAGGAGA</b>	TGAAAGAATA	5002
30	TTACCACATG	AGGAGGAGGG	GCAGTTGCAA	CAGAACACTT	TGGTCTTCCT	ACACCAAGTC	5062
	TGTGAGGGCA	TCCAAGACTG	AATGAAAGCG	CTTTTCTTAT	GCATACAATG	TGAGCAAGAA	5122
	CAAGAACTGT	TTAAGGCACC	TCTGTTCCCA	GCCACTGAAG	AGAGACGTCA	GAAGATGTTA	5182
	GAATAGGTCA	AAACCAAGGC	TCTGGTGGAC	TGAGGGAAGG	TITGTAGCTG	CGTTTAGTGG	5242
	TATACATCTT	TAGTCCCAGC	ATAGGCAGGT	GAATCTCGAG	TTTGAAGCTA	GCCTGGTCTA	5302
<i>3</i> 5	AAAAGGAAGT	TCCAAGACTG	CCAGGGCCAC	ACAGAGGAAA	AAAAAAAACC	CTCTAGAAAA	5362
	ACAAAAATGA	AGACAGGTTC	TCATGTATCG	TAGATTGGCC	TTTAAGTCAC	TTTACCAAGG	5422
	ATGATCTTTG	AACTCCTGAG	TACAGACTGC	GGGTGTGTGC	TACCATGCTT	TATGTGGCCC	5482
	TGGGTTCAAA	CACAGCCCTT	CATATGTATA	TAGCCAAACA	CTCTACAACT	GAGCTACATC	5542
			GTTTTTTGGA				5602
40	TGCAAAGCCA	TTCCTGACCT	GTAAACCTCA	GCTCTCCATC	TCTATAAGAG	GTATAGCCTG	5662
	GGCTAATACC	GTCCAAGTTA	CAACTCCTTG	CTTGCTTTCT	GTTCCTTCTA	GCCTTGGTGA	5722
			CCCCCTCTCT				5782
						CCTGCAACAC	5842
						ACAAAGTGGA	5902
45						TGTAATGAAC	5962
						GTACCCCGTC	6022
						CAAAGTCATG	6082
						GCAAATGTGA	6142
						ATGTCCTCCT	6202
50						ATGGTGCCTT	6262
	GTTTTTTGTT			CITGITTGTA	TTTAATTAAA	ACAAATTGTC	6322
	ATGAGGAAAA	AAAAAAAAA	AAA				6345

## 55 SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 920 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

5

10

15

20

25

30

*3*5

40

45

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Pro Gly Arg Ala Gly Val Ala Arg Phe Cys Leu Leu Ala Leu Ala Leu Gln Leu His Trp Pro Leu Ala Ala Cys Glu Pro Gly Trp Thr Thr Arg Gly Ser Gln Glu Gly Ser Pro Pro Leu Gln His Glu Leu Ile Ile Pro Gln Trp Arg Thr Ser Glu Ser Pro Gly Arg Gly Lys His Pro Leu Arg Ala Glu Leu Arg Val Met Ala Glu Gly Arg Glu Leu Ile Leu Asp Leu Glu Lys Asn Glu His Leu Phe Ala Pro Ala Tyr Thr Glu Thr Cys Tyr Thr Ala Ser Gly Asn Pro Gln Thr Ser Thr Leu Lys Ser Glu Asp His Cys Phe Tyr His Gly Thr Val Arg Asp Val Asp Glu Ser Ser Val Thr Leu Ser Thr Cys Arg Gly Ile Arg Gly Leu Ile Ile Val Arg Ser Asn Leu Ser Tyr Ile Ile Glu Pro Val Pro Asn Ser Asp Ser Gln His Arg Ile Tyr Arg Ser Glu His Leu Thr Leu Pro Pro Gly Asn Cys Gly Phe Glu His Ser Gly Pro Thr Ser Lys Asp Trp Ala Leu Gln Phe Thr His Gln Thr Lys Lys Gln Pro Arg Arg Met Lys Arg Glu Asp Leu His Ser Met Lys Tyr Val Glu Leu Tyr Leu Val Ala Asp Tyr Ala Glu Phe Gln Lys Asn Arg His Asp Gln Asp Ala Thr Lys Arg Lys Leu Met Glu Ile Ala Asn Tyr Val Asp Lys Phe Tyr Arg Ser Leu Asn Ile Arg Ile Ala Leu Val Gly Leu Glu Val Trp Thr His Gly Asp Lys Cys Glu Val Ser Glu Asn Pro Tyr Ser Thr Leu Trp Ser Phe Leu Ser Trp Arg Arg Lys Leu Leu Ala Gln Lys Ser His Asp Asn Ala Gln Leu Ile Thr Gly Arg Ser Phe Gln Gly Thr Thr Ile Gly Leu Ala Pro Leu Met Ala Met Cys Ser Val Tyr Gln Ser Gly Gly Val Ser Met Asp His Ser Glu Asn Ala Ile Gly Val Ala Ser Thr Val Ala His Glu Ile Gly His Asn Phe Gly Met Ser His Asp Ser Ala His Cys Cys Ser Ala Ser Ala Ala Asp Gly Gly Cys Ile Met Ala Ala Ala Thr Gly His Pro Phe Pro Lys Val

P	he	Ser	Trp	Суs	Asn	Arg	Lys	Glu	Leu	Asp	Arg	Tyr	Leu	Gln	Thr	Gly
G	ly	Gly	Met	Сув	Leu	Ser	Asn	Met	Pro	qaA	Thr	Arg	Thr	Leu	Tyr	Gly
G	ly	Arg	Arg	Cys	Gly	Asn	Gly	Tyr	Leu	Glu	qeA	Gly	Glu	Glu	Сув	Ası
C	ув	Gly	Glu	Glu	Glu	Glu	Сув	Lys	neA	Pro	Сув	Cys	цеA	Ala	Ser	Asr
C	ys	Thr	Leu	Lys	Glu	Gly	Ala	Glu	Сув	Ala	His	Gly	Ser	Сув	Cys	His
G	ln	Cys	Lys	Leu	Val	Ala	Pro	Gly	Thr	Gln	Сув	Arg	Glu	Gln	Val	Arg
G	ln	Сув	Asp	Leu	Pro	Glu	Phe	Сув	Thr	Gly	Lys	Ser	Pro	His	Сув	Pro
T	hr	Asn	Tyr	Tyr	Gln	Met	Asp	Gly	Thr	Pro	Сув	Glu	Gly	Gly	Gln	Ala
T	уr	Сув	Tyr	Asn	Gly	Met	Сув	Leu	Thr	Tyr	Gln	Glu	Gln	Cys	Gln	Glr
L	eu	Trp	Gly	Pro	Gly	Ala	Arg	Pro	Ala	Leu	Ąsp	Leu	Сув	Phe	Glu	Arg
ν	al	Asn	Ala	Ala	Gly	Asp	Thr	Tyr	Gly	Asn	Суз	Gly	Lys	Gly	Leu	Asr
g	ly	Gln	Tyr	Arg	Lys	Сув	ser	Pro	Arg	Asp	Ala	Lys	Сув	Xaa	Lys	Ile
G	ln	Сув	Gln	Ser	Thr	Gln	Ala	Arg	Pro	Leu	Glu	Ser	Asn	Ala	Val	Ser
I	le	Asp	Thr	Thr	Ile	Thr	Leu	Asn	Gly	Arg	Arg	Ile	His	Cys	Arg	Gly
T	hr	His	Val	Tyr	Arg	Gly	Pro	Glu	Glu	Glu	Glu	Gly	Glu	Gly	Asp	Met
L	eu	Asp	Pro	Gly	Leu	Val	Met	Thr	Gly	Thr	Lys	Сув	Gly	His	Asn	His
I	le	Cys	Phe	Glu	Gly	Gln	Cys	Arg	neA	Thr	Ser	Phe	Phe	Glu	Thr	Glu
G	ly	Сув	Gly	Lys	Lys	Сув	Asn	Gly	His	Gly	Val	Сув	asa	Asn	neA	Lys
A	sn	Сув	His	Сув	Phe	Pro	Gly	Trp	Ser	Pro	Pro	Phe	Cys	Asn	Thr	Pro
G	ly	qeA	Gly	Gly	Ser	Val	Asp	Ser	Gly	Pro	Leu	Pro	Pro	Lys	Ser	Val
G	ly	Pro	Val	Ile	Ala	Gly	Val	Phe	Ser	Ala	Leu	Phe	Val	Leu	Ala	Val
L	eu	Val	Leu	Leu	Сув	His	Сув	Tyr	Arg	Gln	Şer	His	Lys	Leu	Gly	Lys
P	ro	Ser	Ala	Leu	Pro	Phe	Lys	Leu	Arg	His	Gln	Phe	Ser	Сув	Pro	Phe
A	rg	Val	Ser	Gln	Ser	Gly	Gly	Thr	GJA	His	Ala	Asn	Pro	Thr	Phe	Lys
L	eu	Gln	Thr	Pro	Gln	Gly	Lys	Arg	Lys	Val	Thr	Aen	Thr	Pro	Glu	Sex
L	eu	Arg	Lys	Pro	Ser	His	Pro	Pro	Leu	Arg	Pro	Pro	Pro	Asp	Tyr	Lev
A	rg	Val	Glu	Ser	Pro	Pro	Ala	Pro	Leu	Ser	Ala	His	Leu	Asn	Arg	Ala
A	la	Gly	Ser	Ser	Pro	Glu	Ala	Gly	Ala	Arg	Ile	Glu	Arg	Lys	Glu	Ser

		Ala	Arg	Arg	Pro	Pro	Pro	Ser	Arg	Pro	Met	Pro	Pro	Ala	Pro	Asn	Сув
5		Leu	Leu	Ser	Gln	qaA	Phe	Ser	Arg	Pro	Arg	Pro	Pro	Gln	Lys	Ala	Lev
•		Pro	Ala	Asn	Pro	Val	Pro	Gly	Gln	Arg	Thr	Gly	Pro	Arg	Ser	Gly	Gly
		Thr	Ser	Leu	Leu	Gln	Pro	Pro	Thr	Ser	Gly	Pro	Gln	Pro	Pro	Arg	Pro
10		Pro	Ala	Val	Pro	Val	Pro	Lys	Leu	Pro	Glu	Tyr	Arg	Ser	Gln	Arg	Val
		Gly	Ala	Ile	Ile	Ser	Ser	Lys	Ile 716								
15	SEQ ID	NO:5:															
	<b>~~</b> •					DIOTIC	30.										

20

25

30

35

40

45

50

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3928 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vii) IMMEDIATE SOURCE
- (B)CLONE: JM109(pBSMel γ)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

	GTT	GCAJ	<b>LGGA</b>	TGAC	CGAA	GC G	GAGG	CGGC	G GC	CGCG	CGTI	GAG	CGGA	ACC	TGCC	GAAGC	C 60
	CTC	<b>GCT</b>		GGG Gly												CTA Leu	108
																GGG	156
o																ATT Ile	204
																AGT Ser	252
5																ATT	300
0																GTT Val	348
																GTA Val	396
5																TCC	444
	GCG	GT	r GC	r GTC	agc	GCC	TGC	TTI	r GGJ	A CTC	: AGA	. GGC	TTG	CTG	CAT	TTG	492
o																	

5

	Ala	Val	Ala	Val	ser	Ala	Суз	Phe	Gly	Leu	Arg	Gly	Leu	Leu	His	Leu	
5																TTT Phe	540
																AGA Arg	588
10																GAT Asp	636
																GCT Ala	684
15																AAG Lys	732
20																GAG Glu	780
					GCA Ala												828
25					CTG Leu											CCT Pro	876
																CAG Gln	924
30					TTC Phe												972
35					AAA Lys												1020
					TCA Ser												1068
40					GAG Glu											_	1116
					ATG Met												1164
45					ATC Ile												1212
50					GCG Ala												1260
					CTT Leu												1308
55					AAT Asn												1356

	GGC	ACA	GCG	AAG	GAG	TGT	GAG	GTG	GAC	CCA	TGC	TGT	GAA	GGA	AGC	ACT	1404
	Gly	Thr	Ala	Lys	Glu	Cys	Glu	Val	qaA	Pro	Сув	Сув	Glu	Gly	Ser	Thr	
_																	
5																AAA	1452
	Cys	Lys	Leu	Lys	Ser	Phe	Ala	Glu	Cys	Ala	Tyr	gry	Asp	Cys	Cys	rys	
	CAT	TGC	CAG	TTC	СТТ	CCA	GGA	GGC	TCC	ATG	TGC	AGA	GGG	AAG	ACC	AGT	1500
						Pro											-500
10		-,-									-, -		•				
						GAG											1548
	Glu	Cys	Asp	Val	Pro	Glu	Tyr	Сув	Asn	Gly	Ser	Ser	Gln	Phe	Сув	Pro	
		~~		~~	2 11011	CAG	220	CCA	መኔመ	COTT	TCC	CAC	220	200	222	CCC	1596
15																Ala	1336
	PIO	noy	141			<b>444</b>		<b>4</b> -7	-3-		-,-	<b>42</b>		-01	-1-		
	TAC	TGC	TAC	AAT	GGC	ATG	TGC	CAA	TAT	TAT	GAC	GCG	CAG	TGT	CAG	GTC	1644
	Tyr	Сув	Tyr	Asn	Gly	Met	Сув	Gln	Tyr	Tyr	Asp	Ala	Gln	Сув	Gln	Val	
20								~~~		227		03 M	maa		N. TOVET	CNA	1602
						GCT Ala											1692
	116	FIIC	GLY	361	טעט	~Lu	2,0		<i>_</i>		<b>,</b> 9	p	C ₃ U			<b>524</b>	
•	GTC	AAT	TCT	AAA	GGT	GAC	AGA	TTT	GGC	AAC	TGT	GGT	TTC	TCC	GGC	AGT	1740
25	Val	Asn	Ser	Lys	Gly	qaA	Arg	Phe	Gly	Asn	Cys	Gly	Phe	Ser	Gly	Ser	
•											~~~		-		~~~	<b>63.3</b>	3 700
						Ala										CAA Gln	1788
	GIU	TYL	Dys	Lys	cys	719	****	U.,	70	A_U	204	cy s	O.J	<b></b> , 0	200	<b>J</b> 2	
20	TGC	GAG	AAT	GTA	CAG	GAC	ATG	CCG	GTG	TTT	GGA	ATA	GTA	CCA	GCT	ATC	1836
30	Сув	Glu	Asn	Val	Gln	Asp	Met	Pro	Val	Phe	Gly	Ile	Val	Pro	Ala	Ile	
						CGA	~~~	200		maa	maa	007	CTC.	CAT	marco	CNC	1884
						Arg											1004
		<b>42.11</b>				5	1			-7-		2					
35																AAA	1932
	Leu	Gly	Ser	qaA	Val	Pro	Asp	Pro	Gly	Met	Val	Asn	Glu	Gly	Thr	Lys	
	TOT	CAT	GCT.	ccc	DAG	ATT	TGC	AGG	ТАА	Jelek	CAG	тст	GTA	TAA	GCT	тст	1980
						Ile											
40	-	_															
																GGG	2028
	Val	Leu	neA	Tyr	Asp	Суз	Asp	Ile	Gln	Gly	Lys	Cys	His	GIÀ	H18	GIA	
	GTA	ጥርጥ	AAC	AGC	TAA	AAG	AAT	TGT	CAC	TGT	GAA	GAT	GGC	TGG	GCT	ccc	2076
45						Lys											
		_															
						AAA											2124
	Pro	His	Сув	Asp	Thr	Lys	GIA	Tyr	GTÅ	GIA	Ser	val	Asp	ser	GIY	Pro	
50	ACG	ТДТ	ТАА	GCA	AAG	AGC	ACA	GCA	CTG	AGG	GAC	GGG	CTT	CTG	GTC	TTC	2172
						Ser											
		-			_												
						CCC											2220
55	Phe	Phe	Leu	Ile	Val	Pro	Leu	val	Ala	ALA	ALA	iie	Pne	Leu	rue	116	

					CTA Leu												2268
5					AAT Asn												2316
					CCA Pro												2364
10					TCC Ser												2412
15					GGT Gly												2460
	TTC Phe	CCA Pro	GTA Val	CCA Pro	ACC Thr	TAC Tyr	GCC Ala	GCC Ala	AAG Lys	CAG Gln	CCT Pro	GCG Ala	CAG Gln	TTC Phe	CCG Pro	TCA Ser	2508
20	AGG Arg	CCA Pro	CCT Pro	CCA Pro	CCA Pro	CAA Gln	CCG Pro	AAA Lys	ATA Ile	TCT Ser	TCT Ser	CAG Gln	GGA Gly	AAC Asn	TTG Leu	ATT Ile	2556
	CCG Pro	GCT Ala	CGG Arg	CCC Pro	GCT Ala	CCT Pro	GCA Ala	CCT Pro	CCT Pro	TTA Leu	TAT Tyr	AGC Ser	TCC Ser	CTC Leu	ACC Thr		2601
25	GTT	TTT	TT :	TTT	CTG	T G7	TTT	TTG	AA.	AGCC1	TTC	TCT	CCA	ACC 3	ATGA	ATGTTT ATGAAC AACACA	2661 2721 2781
30	GGAI CTGC TAAT CTGI	ATGTO SCCAT NGATT ACTTY	CA (CATT) OTT)	GCGC GTGGI CAAA1 GACCC	TCCC ATTIF TAAC CAGI	G GC LA TC T G1	GGT( CACT PATTI CATT	TAAI TGAG AGTG1	GTC GTC AAG GTT	SAACO SGATT SCTTT	TTT TAAG TGTC GTT	TTAT ACTA GACO	CGT TCT( ATGC( ATT	rag i Bag ( BCT I I agr	AATG! LATG! AAAC( ACAT(	PTTTCT PTACTG STAATC STATTA	2841 2901 2961 3021
35	TTTI AAGI GTTI	ATATO AACCA ATTGO	GA ( ACA I SCT :	TTA? ATTAC	TAGA	T CA	TAATO ATGAT SATCT	CACI PATACI PTTCI	AAC TTC ACT	Baag( Baaj Bataj	CAGA AGTG ACAA	TATO TGAI TTAT	TCGI ATAT KTADI	AAG ( FGG 1 AGA 1	BAGC PGTGT AATC	CAGCCA ITACAC IACTCA GATTTA	3081 3141 3201 3261
40	GAAC	TGAT	TTC I	AACT(	CACT	T T	TCT TGG1	YATY LDOAT	CAT A GGT	CATO GTT	GTA TAGT	AAGO	ATTY CCA	SCA (	GAG!	rgacag rgttgt aggtgt cttgag	3321 3381 3441 3501
40	AATT	CATO STTTA CAGGO	EAG ( ACA :	CACTI ITTAC SAGAJ	TAAC TAAC AAGGA	T CT	TAAAC TGCT	TCT( TGGGT LAGT(	AAT CCT GTT	TTCI TGTCI	AAAG PCTT PTAG	TTG/	SATGT ACTAI ACTAI	TAD I CAG I	AGTC( PTTT( AATT)	CTCTAG CGTAAA ATACTG	3561 3621 3681
45	TTG	ACACT SGTAT	AA I	AATT!	taati Taadi	C AT	LATT! LTTA:	latti Lgtti	TT	KTADI LAAAI	AATC ATAA	TATT	ATA!	AAG A	AAGT:	ATGATA PTAATA ATATAC AAAAAA	3741 3801 3861 3921 3928
		~~~															

(2) INFORMATION FOR SEQ ID NO:6:

50

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 845 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

5

Met	Gly	Pro	Arg	Ala	Leu	Ser	Pro	Leu	Ala	Ser	Leu	Arg	Leu	Arg	Trp
Leu	Leu	Ala	Сув	Gly	Leu	Leu	Gly	Pro	Val	Leu	Glu	Ala	Gly	Arg	Pro
Asp	Leu	Glu	Gln	Thr	Val	His	Leu	Ser	Ser	Tyr	Glu	Ile	Ile	Thr	Pro
Trp	Arg	Leu	Thr	Arg	Glu	Arg	Arg	Glu	Ala	Leu	Gly	Pro	Ser	Ser	Gln
Gln	Ile	Ser	Tyr	Val	Ile	Gln	Ala	Gln	Gly	Lys	Gln	His	Ile	Ile	His
Leu	Glu	Arg	Asn	Thr	qeA	Leu	Leu	Pro	Asn	Авр	Phe	Val	Val	Tyr	Thr
Tyr	Asp	Lys	Glu	Gly	Ser	Leu	Leu	Ser	Ąsp	His	Pro	Asn	Val	Gln	Ser
Hië	Cys	His	Tyr	Arg	Gly	Tyr	Val	Glu	Gly	Val	Gln	Asn	Ser	Ala	Val
Ala	Val	Ser	Ala	Сув	Phe	Gly	Leu	Arg	Gly	Leu	Leu	His	Leu	Glu	Asn
Ala	Ser	Phe	Gly	Ile	Glu	Pro	Leu	His	Asn	Ser	Ser	His	Phe	Glu	His
Ile	Phe	Tyr	Pro	Met	Asp	Gly	Ile	His	Gln	Glu	Pro	Leu	Arg	Cys	Gly
Val	Ser	Asn	Arg	Asp	Thr	Glu	Lys	Glu	Gly	Thr	Gln	Gly	qaA	Glu	Glu
Glu	His	Pro	Ser	Val	Thr	Gln	Leu	Leu	Arg	Arg		Arg -205		Val	Leu
Pro	Gln	Thr	Arg	Tyr	Val	Glu	Leu	Phe	Ile	Val	Val	qeA	Lys	Glu	Arg
Tyr	Asp	Met	Met	Gly	Arg	Asn	Gln	Thr	Ala	Val	Arg	Glu	Glu	Met	Ile
Arg	Leu	Ala	Asn	Tyr	Leu	Asp	Ser	Met	Tyr	Ile	Met	Leu	Asn	Ile	Arg
Ile	Val	Leu	Val	Glÿ	Leu	Glu	Ile	Trp	Thr	Asp	Arg	Asn	Pro	Ile	Asn
Ile	Ile	Gly	Gly	Ala	Gly	qaA	Val	Leu	Gly	Asn	Phe	Val	Gln	Trp	Arg
			Leu											•	
	_		Gly												
	_		Arg												
			Thr												
	_		Asn												
	-		Met												
			Glu												
Сув	Leu	Leu	Asn	Ile	Pro	Lys	Pro	Asp	Glu	Ala	Tyr	Ser	Ala	Pro	Ser

Сув	Gly	Asn	Lys	Leu	Val	Двр	Pro	Gly	Glu	Glu	Сув	двр	Сув	Gly	Thr
Ala	Lув	Glu	Сув	Glu	Val	qaA	Pro	Cys	Сув	Glu	Gly	Ser	Thr	Cys	Lys
Leu	Lys	Ser	Phe	Ala	Glu	Cys	Ala	тут	Gly	qaA	Cys	Суз	Lys	Авр	Сув
Gln	Phe	Leu	Pro	Gly	Gly	Ser	Met	Cys	Arg	Gly	Lys	Thr	Ser	Glu	Cys
Asp	Val	Pro	Glu	Tyr	Cys	Asn	Gly	Ser	Ser	Gln	Phe	Суз	Pro	Pro	qeA
Val	Phe	Ile	Gln	Asn	Gly	Tyr	Pro	Сув	Gln	Asn	Ser	Lys	Ala	Tyr	Сув
Tyr	Asn	Gly	Met	Сув	Gln	Tyr	Tyr	Asp	Ala	Gln	Сув	Gln	Val	Ile	Phe
Gly	Ser	Lys	Ala	Lys	Ala	Ala	Pro	Arg	Asp	Cys	Phe	Ile	Glu	Val	Asn
Ser	Lys	Gly	Asp	Arg	Phe	Gly	Asn	Cys	Gly	Phe	Ser	Gly	Ser	Glu	Tyr
Lys	Lys	Сув	Ala	Thr	Gly	Asn	Ala	Leu	Cys	Gly	Lys	Leu	Gln	Сув	Glu
Asn	Val	Gln	Asp	Met	Pro	Val	Phe	Gly	Ile	Val	Pro	Ala	Ile	Ile	Gln
Thr	Pro	Ser	Arg	Gly	Thr	Lys	Сув	Trp	Gly	Val	Asp	Phe	Gln	Leu	Gly
Ser	qaA	Val	Pro	Asp	Pro	Gly	Met	Val	Asn	Glu	Gly	Thr	Lys	Сув	Asp
Ala	Gly	Lys	Ile	Сув	Arg	Asn	Phe	Gln	Суз	Val	Asn	Ala	Ser	Val	Leu
Asn	Tyr	Asp	Сув	Asp	Ile	Gln	Gly	Lys	Сув	His	Gly	His	Gly	Val	Сув
Asn	Ser	Asn	Lys	Asn	Сув	His	Сув	Glu	Asp	Gly	Trp	Ala	Pro	Pro	His
Сув	Asp	Thr	Lys	Gly	Tyr	Gly	Gly	Ser	Val	qaA	Ser	Gly	Pro	Thr	Tyr
Asn	Ala	Lys	Ser	Thr	Ala	Leu	Arg	qaA	Gly	Leu	Leu	Val	Phe	Phe	Phe
Leu	Ile	Val	Pro	Leu	Val	Ala	Ala	Ala	Ile	Phe	Leu	Phe	Ile	Lys	Arg
Asp	Glu	Leu	Arg	Lys	Thr	Phe	Arg	Lys	Lys	Arg	Ser	Gln	Met	Ser	Asp
Gly	Arg	Asn	Gln	Ala	Asn	Val	Ser	Arg	Gln	Pro	Gly	qaA	Pro	Ser	Ile
Ser	Arg	Pro	Pro	Gly	Gly	Pro	Asn	Val	Ser	Arg	Pro	Pro	Gly	Gly	Pro
Gly	Val	Ser	Arg	Pro	Pro	Gly	Gly	Pro	Gly	Val	Ser	Arg	Pro	Pro	Gly
Gly	Pro	Gly	Val	Ser	Arg	Pro	Pro	Pro	Gly	His	Gly	Asn	Arg	Phe	Pro
Val	Pro	Thr	Tyr	Ala	Ala	Lys	Gln	Pro	Ala	Gln	Phe	Pro	Ser	Arg	Pro
Pro	Pro	Pro	Gln	Pro	Lys	Ile	Ser	Ser	Gln	Gly	Asn	Leu	Ile	Pro	Ala
Arg	Pro	Ala	Pro	Ala	Pro	Pro	Leu	Tyr	Ser	Ser	Leu	Thr 640			

SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

55

5		(E (C) LEN () TYP () STR () TOP	E: nuc ANDE	leic ac	id S: not		ant						
10		(ii) MC (iii) HY (iv) AN (vii) IN	POTH	HETICA ENSE: ATE S	AL: NO NO SOURC	D DE	ıMa20	١٥١						
15		(xi) SE						·	NO:7:					
20		CCT Pro												46
		GAG Glu												96
25		CAC His												144
		ATC Ile												192
30		GCT Ala												240
a-		GAA Glu												286
35		TGC Cys								_				321
40	SEC) ID N	D:8:											
		(i) SE(QUEN	CE CH	IARAC	CTERI	STICS	S :			,			
45		(B) LEN) TYP) TOP	E: ami	no aci	d	cids							
		(ii) MC						O 15 :	10.5					
50		(XI) SE	:UUE	NCE D	ESCH	iiP110	IN: SE	Q ID N	NO:8:					

		Lys 1	Pro	Ala	Gly	Thr 5	Ala	Сув	Arg	Asp	Ser 10	Ser	Asn	Ser	Сув	Asp 15	Leu
5			Glu	Phe	Cys 20		Gly	Ala	Ser	Pro 25		Сув	Pro	Ala	neA 30		Tyr
•		Leu	His	Asp 35		His	Ser	Сув	Gln 40	Asp	Val	Asp	Gly	Tyr 45		Xaa	Asn
		Gly	Ile 50		Gln	Thr	His	Glu 55		Gln	Сув	Val	Thr 60	Leu	Trp	Gly	Pro
10		Gly 65		Lys	Pro	Ala	Pro		Ile	Cys	Phe	Glu 75		Val	Asn	Ser	Ala 80
15		Gly	Glu	Pro	Tyr	Gly 85		Cys	Gly	' Lys	Va) 90		Lys	Ser	Ser	Phe 95	
		Lys	Сув	Glu	Met 100	Arg		Ala	Lys	Cys 105	Gly		1			,,,	
20																	
	SEQ ID	NO:9:															
	(i) S	EQUE	NCE	CHAR	ACTE	RISTIC	S:										
25		(A) LE (B) TY				oairs											
		(C) ST				ot rele	vant										
		(D) TC	POLO	GY: I	inear												
30		MOLEC															
		HYPO ⁻ ANTI-S			NO												
	(vii)	IMME	DIATE	SOU	RCE												
35		(B) CL	ONE:	JM10	9 (pBS	shuMγ	G238))									
	(xi)	SEQU	ENCE	DESC	CRIPTI	ON: S	EQ ID	NO:9	:								
40																	
45																	
50																	
50																	
55																	

	GCA	AAG	AGC	TGC	ATC	ATG	AAT	TCA	GGA	GCA	TCG	GGT	TCC	AGA	AAC	TTT	48
	Ala	Lys	Ser	Сув	Ile	Met	Asn	Ser	Gly	Ala	Ser	Gly	Ser	Arg	Asn	Phe	
5	AGC	AGT	TGC	AGT	GCA	GAG	GAC	TTT	GAG	AAG	TTA	ACT	TTA	AAT	AAA	GGA	96
																Gly	
																GCT	144
10	Gly	Asn	Сув	Leu	Leu	neA	Ile	Pro	Lys	Pro	qaA	Glu	Ala	Tyr	Ser	Ala	
	CCC	TCC	TGT	GGT	AAT	AAG	TTG	GTG	GAC	GCT	GGG	GAA	GAG	TGT	GAC	TGT	192
	Pro	Ser	Сув	Gly	Asn	Lys	Leu	Val	Ąsp	Ala	Gly	Glu	Glu	Сув	Asp	Сув	
	GGT	ACT	CCA	AAG	GAA	TGT	GAA	TTG	GAC	ССТ	TGC	TGC	GAA	GGA	AGT	ACC	. 240
15	Gly	Thr	Pro	Lys	Glu	Сув	Glu	Leu	qaA	Pro	Сув	Сув	Glu	Gly	Ser	Thr	
	TGT	AAG	CTT	AAA	TCA	TTT	GCT	GAG	TGT	GCA	TAT	GGT	GAC	TGT	TGT	AAA	288
																Lys	
20	GAC	TGT	CGG	TTC	CTT	CCA	GGA	GGT	ACT	TTA	TGC	CGA	GGA	AAA	ACC	AGT	336
																Ser	
																CAG	384
25	Glu	Суз	Asp	Val	Pro	Glu	Tyr	Сув	Asn	Gly	Ser	Ser	Gln	Phe	Cys	Gln	
	CCA	GAT	GTT	TTT	ATT	CAG	AAT	GGA	TAT	CCT	TGC	CAG	AAT	AAC	AAA	GCC	432
	Pro	Asp	Val	Phe	Ile	Gln	Asn	Gly	Tyr	Pro	Cys	Gln	Asn	Asn	Lys	Ala	
~~																GTC	480
30	Tyr	Cys	Tyr	Asn	Gly	Met	СЛа	Gln	Tyr	Tyr	Asp	Ala	Gln	Сув	Gln	Val	
																GAA	528
	Ile	Phe	Gly	Ser	Lys	Ala	Lys	Ala	Ala	Pro	Lys	Asp	Cys	Phe	Ile	Glu	
35						GAC											576
	Val	Asn	Ser	Lys	Gly	Asp	Arg	Phe	Gly	Asn	Сув	Gly	Phe	Ser	Gly	Asn	

	GAA	TAC	AAG	aag	TGT	GCC	ACT	GGG	AAT	GCT	TTG	TGT	GGA	AAG	CTT	CAG	624
	Glu	Tyr	Lys	Lys	Cys	Ala	Thr	Gly	Asn	Ala	Leu	САв	Gly	Lys	Leu	Gln	
5	TGT	GAG	AAT	GTA	CAA	GAG	ATA	CCT	GTA	TTT	GGA	ATT	GTG	CCT	GCT	ATT	672
	CAa	Glu	Asn	Val	Gln	Glu	Ile	Pro	Val	Phe	Gly	Ile	Val	Pro	Ala	Ile	
	ATT	CAA	ACG	CCT	AGT	CGA	GGC	ACC	AAA	TGT	TGG	GGT	GTG	GAT	TTC	CAG	720
10	Ile	Gln	Thr	Pro	Ser	Arg	Gly	Thr	Lys	Сув	Trp	Gly	Val	Asp	Phe	Gln	
	CTA	GGA	TCA	GAT	GTT	CCA	GAT	ССТ	GGG	ATG	GTT	AAC	GAA	GGC	ACA	AAA	768
	Leu	Gly	Ser	Asp	Val	Pro	Asp	Pro	Gly	Met	Val	aea	Glu	Gly	Thr	Lys	
15	-															TCT	816
.•	Сув	Gly	Ala	Gly	Lys	Ile	Сув	Arg	Asn	Phe	Gln	Сув	Val	Asp	Ala	Ser	
																GGG	864
	Val	Leu	Asn	Tyr	qaA	Сув	Asp	Val	Gln	Lys	Lys	Сув	His	Gly	His	Gly	
20						AAG											912
	Val	Сув	Asn	Ser	Asn	Lys	Asn	Сув	His	Сув	Glu	Asn	Gly	Trp	Leu	Pro	
	-															TCG	960
25	Gln	Ile	Val	Arg	Leu	Lys	qaA	Thr	Arg	ser	Ser	Leu	ser	Ile	Pro	ser	
	ACC		A				•										967
	Thr	ser															

30 (2) INFORMATION FOR SEQ ID NO:10:

*3*5

40

45

50

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 322 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

		Mla	Lva	Ser	Cvs	Ile	Met	Asn	Ser	Glv	Ala	Ser	Glv	Ser	Ara	Asn	Phe
		1	_		_	5					10					15	
5		Ser	Ser	Суз	Ser 20	Ala	Glu	Asp	Phe	G1u 25	Lys	Leu	Thr	Leu	Asn 30	Lys	GIA
		Gly	Asn	Сув 35	Leu	Leu	Asn	Ile	Pro 40	Lys	Pro	qaA	Glu	Ala 45	Tyr	Ser	Ala
		Pro	Ser 50	Сув	Gly	Asn	Lys	Leu 55	Val	Asp	Ala	Gly	Glu 60	Glu	Сув	Двр	Cys
10		Gly 65		Pro	Lys	Glu	Сув 70	-	Leu	Asp	Pro	Cys 75	Сув	Glu	Gly	Ser	Thr 80
		_	Lys	Leu	Lys	Ser 85		Ala	Glu	Сув	Ala 90		Gly	Asp	Сув	Cys 95	
15		qaA	Сув	Arg	Phe 100		Pro	Gly	Gly	Thr 105		Сув	Arg	Gly	Lys 110		Ser
		Glu	Сув	_	Val	Pro	Glu	Tyr	Cys 120	-	Gly	Ser	Ser	Gln 125		Cys	Gln
		Pro	Asp	115 Val	Phe	Ile	Gln	Asn		Tyr	Pro	Суз	Gln		Asn	Lys	Ala
20																	
			130					135					140				
		Tyr		Tyr	Asn	Gly	Met		Gln	Tyr	Tyr	Asp	Ala	Gln	Сув	Gln	Val
		145	_				150					155					160
25				_	Ser	165		•			170					175	
		Val	Asn	Ser	Lys 180	Gly	qeA	Arg	Phe	Gly 185	neA	Сув	Gly	Phe	Ser 190	Gly	Asn
30		Glu	Tyr	Lys 195	Lys	Сув	Ala	Thr	Gly 200	Asn	Ala	Leu	Сув	Gly 205	Lys	Leu	Gln
		Cys	Glu 210	neA	Val	Gln	Glu	Ile 215	Pro	Val	Phe	Gly	11e 220	Val	Pro	Ala	Ile
		Ile	Gln	Thr	Pro	Ser	Arg	Gly	Thr	Lys	Сув	Trp	Gly	Val	Asp	Phe	Gln
25		225					230			_		235		_			240
35		Leu	Gly	Ser	Asp	Val 245	Pro	Asp	Pro	Gly	Met 250	Val	Asn	Glu	Gly	Thr 255	Lys
		Сув	Gly	Ala	Gly 260	Lys	Ile	Сув	Arg	Asn 265	Phe	Gln	Сув	Val	Авр 270	Ala	Ser
40		Val	Leu	Asn 275	Tyr	Asp	Сув	Asp	Val 280	Gln	Lys	Lys	Cys	His 285	Gly	His	Gly
		Val	_	Asn	Ser	Asn	Lys		Сув	His	CAa	Glu		Gly	Trp	Leu	Pro
		Gln	290 Ile	Val	Arg	Leu	Lvs	295 Asp	Thr	Arq	Ser	Ser	300 Leu	Ser	Ile	Pro	Ser
45		305			3		310			•		315					320
45		Thr	Ser														
	SEQ ID	NO:11	:														
50	(i) S	EQUE	NCE	CHAR	ACTE	RISTIC	S:										
		(A) I F	NGTH	ł: 2848	3 base	pairs											
		• •		ucleic		,											
55		(C) S1	[RAN		ESS: n	ot rele	vant										

(ii) MOLECULE TYPE: cDNA (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO	
(vii) IMMEDIATE SOURCE	

(B) CLONE:JM109(pMeI α -25C) JM109 (pMeI α -26N)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

10	GGG	GAC	CTC	TGG	ATC	CCA	GTG	AAG	AGC	TTC	GAC	TCC	AAG	AAT	CAT	CCA	48
	Gly	Asp	Leu	Trp	Ile	Pro	Val	Lys	Ser	Phe	Asp	Ser	Lys	Asn	His	Pro	
	GAA	GTG	CTG	AAT	ATT	CGA	CTA	CAA	CGG	GAA	AGC	AAA	GAA	CTG	ATC	ATA	96
	Glu	Val	Leu	Asn	Ile	Arg	Leu	Gln	Arg	Glu	Ser	Lys	Glu	Leu	Ile	Ile	
15	220		GAA	nca.	AAT	GNA	CCT	כיויר	ATT	acc	AGC	AGT	TTC	ACG	GAA	ACC	144
			Glu														
	CAC	тат	CTG	CAA	GAC	GGT	ACT	GAT	GTC	TCC	CTC	GCT	CGA	AAT	TAC	ACG	192
20			Leu														
	GGT	CAC	TGT	TAC	TAC	CAT	GGA	CAT	GTA	CGG	GGA	TAT	TCT	GAT	TCA	GCA	240
			Cvs													_	

5

	GTC	AGT	CTC	AGC	ACG	TGT	TCT	GGT	CTC	AGG	GGA	CTT	ATT	GGG	TTT	GAA	288
	Val	Ser	Leu	Ser	Thr	Сув	Ser	Gly	Leu	Arg	Gly	Leu	Ile	Gly	Phe	Glu	
5																	
3																TAC	336
	MDM	GIU	361	-7-	*44	bea	024			275		744		AU.	~~ y	-1-	
																GGA	384
10	Lys	Leu	Phe	Pro	Ala	Lys	Lys	Leu	Lys	Ser	Val	Arg	Gly	Ser	Сув	Gly	
,•	ም ረን	CAT	CNC	AAC	aca.	CCA	AAC	CTC	CCT	GCA	AAG	אמ דממ	OTC.	TTT	CCA	CCA	432
		His															432
											_						
15																GCA	480
	Pro	Ser	GIN	Thr	Trp	ATA	Arg	Arg	HIB	гув	Arg	GIU	THE	ren	ГÀВ	Ala	
	ACT	AAG	TAT	GTG	GAG	CTG	GTG	ATC	GTG	GCA	GAC	AAC	CGA	GAG	TTT	CAG	528
	Thr	Lys	Tyr	Val	Glu	Leu	Val	Ile	Val	Ala	qaA	Asn	Arg	Glu	Phe	Gln	
20		~~~	~~		63.00	~	CAA	222	Circui	220	CDG	CCA	- Arrest	אדא	CNG	ATT	576
		Gln															376
				-7-				_, _		_, _							
																GTG	624
25	Ala	Asn	His	Val	Asp	Lys	Phe	Tyr	Arg	Pro	Leu	Asn	Ile	Arg	Ile	Val	
	TTG	GTA	GGC	GTG	GAA	GTG	TGG	AAT	GAC	ATG	GAC	AAA	TGC	TCT	GTA	AGT	672
	Leu	Val	Gly	Val	Glu	Val	Trp	Asn	qaA	Met	qeA	Lys	Cys	Ser	Val	Ser	
	~~~		<b>553</b>	<b>650</b>		100	CTIC	CAM	CAA	Trees		CAC	TCC	200	224	ATG	720
30		Asp															720
		_														•	
		CTT															768
	Lys	Leu	Leu	Pro	Arg	Lys	ser	HIS	Авр	ASN	Ala	GIN	Leu	vai	ser	Gly	
35	GTT	TAT	TTC	CAA	GGG	ACC	ACC	ATC	GGC	ATG	GCC	CCA	ATC	ATG	AGC	ATG	816
	Val	Tyr	Phe	Gln	Gly	Thr	Thr	Ile	Gly	Met	Ala	Pro	Ile	Met	Ser	Met	
	mcc	ACG	CON	CNC	CRC	ur Carr	aaa	CCN	אדיר	CTC	איניים	CNC	ሮአሞ	TCA	GNC	አአጥ	864
		Thr															004
40	-,-			•			-	-				-			-		
		CTT															912
	Pro	Leu	GIĄ	ATA	ATS	vaı	Thr	Leu	Ala	H18	GIU	ren	GIA	HIB	ASN	rne	
	GGG	ATG	AAT	CAT	GAC	ACA	CTG	GAC	AGG	GGC	TGT	AGC	TGT	CAA	ATG	GCG	960
45	Gly	Met	Asn	His	qaA	Thr	Leu	qsA	Arg	Gly	Сув	Ser	Суз	Gln	Met	Ala	
	C-CT-CTT	GAG		CCN	ccc	TYCE	እምሮ	בעדת	A D C	CCT	TCC	200	ccc	ጥልሮ	CCA	ملحلمك	1008
		Glu															1000
			•	•	_												
50		ATG															1056
	Pro	Met	Val	Pne	ser	ser	сув	ser	arg	rys	Asp	ren	GIU	Tnr	ser	Ten	
	GAG	AAA	GGA	ATG	GGG	GTG	TGC	CTG	TTT	AAC	CTG	CCG	GAA	GTC	AGG	GAG	1104
ee		Lys															
55																	

																GAG Glu	1152
5				TGT Cys													1200
				TGT Cys													1248
10				GAC Asp													1296
15	TCC Ser	AGC Ser	AAC Asn	TCC Ser	TGT Cys	GAC Asp	CTC Leu	CCA Pro	GAG Glu	TTC Phe	TGC Cys	ACA Thr	GGG Gly	GCC Ala	AGC Ser	CCT Pro	1344
				GCC Ala													1392
20	Val	Asp	Gly	TAC Tyr	Суя	Tyr	Asn	Gly	Ile	Сув	Gln	Thr	His	Glu	Gln	Gln	1440
25	Сув	Val	Thr	CTC Leu	Trp	Gly	Pro	Gly	Ala	Lys	Pro	Ala	Pro	Gly	Ile	Сув	1488
	Phe	Glu	Arg	GTC Val	Asn	Ser	Ala	Gly	Asp	Pro	Tyr	Gly	Asn	Сув	Gly	Lys	1536
30	Val	Ser	Lys	AGT Ser	Ser	Phe	Ala	Lys	Сув	Glu	Met	Arg	qaA	Ala	Lys	Сув	1584
	Gly	Lys	Ile	Gln	Сув	Gln	Gly	Gly	Ala	Ser	Arg	Pro	Val	Ile	Gly		1632
35	Asn	Ala	Val	Ser	Ile	Glu	Thr	Asn	Ile	Pro	Leu	Gln	Gln	Gly	Gly		1728
40	Ile	Leu	Сув	Arg	Gly	Thr	His	Val	Tyr	Leu	Gly	Asp	qaA	Met	Pro		1776
	Pro	Gly	Leu	Val CAA	Leu	Ala	Gly	Thr	Lys	Сув	Ala	qeA	Gly	Lys	Ile	Сув	1824
45	Leu	Asn	Arg	Gln	Сув	Gln	Asn	Ile	Ser	Val	Phe	Gly	Val	His	Glu	Cys	1872
50	Ala	Met	Gln	Cys	His	Gly	Arg	Gly	Val	Сув	Asn	Asn	Arg	Lys	Asn	Сув	1920
JU	GGA	GGA	AGC	Ala ACA	GAC	AGC	GGC	ccc	ATC	CGG	CAA	GCA	GAA	GCA	AGG	CAG	1968
55	_	_		Thr GAG													2016

# Glu Ala Ala Glu Ser Asn Arg Glu Arg Gly Gln Gly Gln Glu Pro Val

5	GGA TCG CAG GAG CAT GCG TCT ACT GCC TCA CTG ACA CTC ATC TGA Gly Ser Gln Glu His Ala Ser Thr Ala Ser Leu Thr Leu Ile *	2061
	GCCCTCCCAT GACATGGAGA CCGTGACCAG TGCTGCTGCA GAGGAGGTCA CGCGTCCCCA AGGCCTCCTG TGACTGGCAG CATTGACTCT GTGGCTTTGC CATCGTTTCC ATGACAACAG	2121 2181
	ACACAACACA GTTCTCGGGG CTCAGGAGGG GAAGTCCAGC CTACCAGGCA CGTCTGCAGA	2241
10	AACAGTGCAA GGAAGGGCAG CGACTTCCTG GTTGAGCTTC TGCTAAAACA TGGACATGCT	2301
	TCAGTGCTGC TCCTGAGAGA GTAGCAGGTT ACCACTCTGG CAGGCCCCAG CCCTGCAGCA	2361
	AGGAGGAAGA GGACTCAAAA GTCTGGCCTT TCACTGAGCC CCCACAGCAG TGGGGGAGAA	2421
	GCAAGGGTTG GGCCCAGTGT CCCCTTTCCC CAGTGACACC TCAGCCTTGG CAGCCCTGAT	2481
	GACTGGTCTC TGGCTGCAAC TTAATGCTCT GATATGGCTT TTAGCATTTA TTATATGAAA	2541
15	ATAGCAGGGT TITAGTTTTT AATTTATCAG AGACCCTGCC ACCCATTCCA TCTCCATCCA	2601
	AGCAAACTGA ATGGCATTGA AACAAACTGG AGAAGAAGGT AGGAGAAAGG GCGGTGAACT	2661
	CTGGCTCTTT GCTGTGGACA TGCGTGACCA GCAGTACTCA GGTTTGAGGG TTTGCAGAAA	2721
	GCCAGGGAAC CCACAGAGTC ACCAACCCTT CATTTAACAA GTAAGAATGT TAAAAAGTGA	2781
	AAACAATGTA AGAGCCTAAC TCCATCCCCC GTGGCCATTA CTGCATAAAA TAGAGTGCAT	2841
20	CCCGCCC	2848

#### SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 686 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30

35

40

45

50

55

25

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Gly Asp Leu Trp Ile Pro Val Lys Ser Phe Asp Ser Lys Asn His Pro Glu Val Leu Asn Ile Arg Leu Gln Arg Glu Ser Lys Glu Leu Ile Ile Asn Leu Glu Arg Asn Glu Gly Leu Ile Ala Ser Ser Phe Thr Glu Thr His Tyr Leu Gln Asp Gly Thr Asp Val Ser Leu Ala Arg Asn Tyr Thr Gly His Cys Tyr Tyr His Gly His Val Arg Gly Tyr Ser Asp Ser Ala Val Ser Leu Ser Thr Cys Ser Gly Leu Arg Gly Leu Ile Gly Phe Glu Asn Glu Ser Tyr Val Leu Glu Pro Met Lys Ser Ala Thr Asn Arg Tyr Lys Leu Phe Pro Ala Lys Lys Leu Lys Ser Val Arg Gly Ser Cys Gly Ser His His Asn Thr Pro Asn Leu Ala Ala Lys Asn Val Phe Pro Pro Pro Ser Gln Thr Trp Ala Arg Arg His Lys Arg Glu Thr Leu Lys Ala Thr Lys Tyr Val Glu Leu Val Ile Val Ala Asp Asn Arg Glu Phe Gln Arg Gln Gly Lys Asp Leu Glu Lys Val Lys Gln Arg Leu Ile Glu Ile Ala Asn His Val Asp Lys Phe Tyr Arg Pro Leu Asn Ile Arg Ile Val

**5**.

10

15

20

25

30

*3*5

40

45

50

55

Leu Val Gly Val Glu Val Trp Asn Asp Met Asp Lys Cys Ser Val Ser Gln Asp Pro Phe Thr Ser Leu His Glu Phe Leu Asp Trp Arg Lys Met Lys Leu Leu Pro Arg Lys Ser His Asp Asn Ala Gln Leu Val Ser Gly Val Tyr Phe Gln Gly Thr Thr Ile Gly Met Ala Pro Ile Met Ser Met Cys Thr Ala Asp Gln Ser Gly Gly Ile Val Met Asp His Ser Asp Asn Pro Leu Gly Ala Ala Val Thr Leu Ala His Glu Leu Gly His Asn Phe Gly Met Asn His Asp Thr Leu Asp Arg Gly Cys Ser Cys Gln Met Ala Val Glu Lys Gly Gly Cys Ile Met Asn Ala Ser Thr Gly Tyr Pro Phe Pro Met Val Phe Ser Ser Cys Ser Arg Lys Asp Leu Glu Thr Ser Leu Glu Lys Gly Met Gly Val Cys Leu Phe Asn Leu Pro Glu Val Arg Glu Ser Phe Gly Gly Gln Lys Cys Gly Asn Arg Phe Val Glu Glu Gly Glu Glu Cys Asp Cys Gly Glu Pro Glu Glu Cys Met Asn Arq Cys Cys Asn Ala Thr Thr Cys Thr Leu Lys Pro Asp Ala Val Cys Ala His Gly Leu Cys Cys Glu Asp Cys Gln Leu Lys Pro Ala Gly Thr Ala Cys Arg Asp Ser Ser Asn Ser Cys Asp Leu Pro Glu Phe Cys Thr Gly Ala Ser Pro His Cys Pro Ala Asn Val Tyr Leu His Asp Gly His Ser Cys Gln Asp Val Asp Gly Tyr Cys Tyr Asn Gly Ile Cys Gln Thr His Glu Gln Gln Cys Val Thr Leu Trp Gly Pro Gly Ala Lys Pro Ala Pro Gly Ile Cys Phe Glu Arg Val Asn Ser Ala Gly Asp Pro Tyr Gly Asn Cys Gly Lys Val Ser Lys Ser Ser Phe Ala Lys Cys Glu Met Arg Asp Ala Lys Cys Gly Lys Ile Gln Cys Gln Gly Gly Ala Ser Arg Pro Val Ile Gly Thr Asn Ala Val Ser Ile Glu Thr Asn Ile Pro Leu Gln Gln Gly Gly Arg Ile Leu Cys Arg Gly Thr His Val Tyr Leu Gly Asp Asp Met Pro Asp Pro Gly Leu Val Leu Ala Gly Thr Lys Cys Ala Asp Gly Lys Ile Cys Leu Asn Arg Gln Cys Gln Asn Ile Ser Val Phe Gly Val His Glu Cys Ala Met Gln Cys His Gly Arg Gly Val Cys Asn Asn Arg Lys Asn Cys His Cys Glu Ala His Trp Ala Pro Pro Phe Cys Asp Lys Phe Gly Phe

Gly Gly Ser Thr Asp Ser Gly Pro Ile Arg Gln Ala Glu Ala Arg Gln Glu Ala Ala Glu Ser Asn Arg Glu Arg Gly Gln Gly Gln Glu Pro Val Gly Ser Gln Glu His Ala Ser Thr Ala Ser Leu Thr Leu Ile

SEQ	ID	NO:13:	

	SEQ	טו	NO	:13
10				

5

15

20

(i) :	SEQI	JENCE	CHARACT	ERIST	ICS:
-------	------	-------	---------	-------	------

- (A) LENGTH: 394 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

	فافافا	GAA	GAG	161	GAT	161	GGA	GAA	GAA	GAG	GAA	161	AAC	AAC	CCC	TGC	48
25	Gly	Glu	Glu	Cys	Asp	Сув	Gly	Glu	Glu	Glu	Glu	Сув	Asn	Asn	Pro	Сув	
	TGC	AAT	GCC	TCT	AAT	TGT	ACC	CTG	AGG	CCG	GGG	GCG	GAG	TGT	GCT	CAC	96
	Cys	Asn	Ala	Ser	Asn	Сув	Thr	Leu	Arg	Pro	Gly	Ala	Glu	Сув	Ala	His	
30	GGC	TCC	TGC	TGC	CAC	CAG	TGT	AAG	CTG	TTG	GCT	CCT	GGG	ACC	CTG	TGC	144
	Gly	Ser	Сув	Cys	His	Gln	Сув	Lys	Leu	Leu	Ala	Pro	Gly	Thr	Leu	Cys	
	CGC	GAG	CAG	GCC	AGG	CAG	TGT	GAC	CTC	CCG	GAG	TTC	TGT	ACG	GGC	AAG	192
	Arg	Glu	Gln	Ala	Arg	Gln	Cys	qeA	Leu	Pro	Glu	Phe	Сув	Thr	Gly	Lys	
35	TCT	CCC	CAC	TGC	CCT	ACC	AAC	TTC	TAC	CAG	ATG	GAT	GGT	ACC	ccc	TGT	240
	Ser	Pro	His	Cys	Pro	Thr	Asn	Phe	Tyr	Gln	Met	qeA	Gly	Thr	Pro	Сув	
	GAG	GGC	GGC	CAG	GCC	TAC	TGC	TAC	AAC	GGC	ATG	TGC	CTC	ACC	TAC	CAG	288
40	Glu	Gly	Gly	Gln	Ala	Tyr	Cys	Tyr	Asn	Gly	Met	Сув	Leu	Thr	Tyr	Gln	
	GAG	CAG	TGC	CAG	CAG	CTG	TGG	GGA	CCC	GGA	GCC	CGA	CCT	GCC	CCT	GAC	336
	Glu	Gln	Сув	Gln	Gln	Leu	Trp	Gly	Pro	Gly	Ala	Arg	Pro	Ala	Pro	qeA	
45	CTC	TGC	TTC	GAG	AAG	GTG	AAT	GTG	GCA	GGA	GAC	ACC	TTT	GGA	AAC	TGT	384
	Leu	Сув	Phe	Glu	Lys	Val	Asn	Val	Ala	Gly	qaA	Thr	Phe	Gly	neA	Сув	
	GGA	AAG	GAC	A													394
	Gly	Lys	Asp														

SEQ ID NO:14:

50

55

#### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 131 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:
- 5 Gly Glu Glu Cys Asp Cys Gly Glu Glu Glu Glu Cys Asn Asn Pro Cys 1 Cys Asn Ala Ser Asn Cys Thr Leu Arg Pro Gly Ala Glu Cys Ala His Gly Ser Cys Cys His Gln Cys Lys Leu Leu Ala Pro Gly Thr Leu Cys 10 Arg Glu Gln Ala Arg Gln Cys Asp Leu Pro Glu Phe Cys Thr Gly Lys Ser Pro His Cys Pro Thr Asn Phe Tyr Gln Met Asp Gly Thr Pro Cys 65 70 75 15 Glu Gly Gly Gln Ala Tyr Cys Tyr Asn Gly Met Cys Leu Thr Tyr Gln 85 90 Glu Gln Cys Gln Gln Leu Trp Gly Pro Gly Ala Arg Pro Ala Pro Asp 105 110 Leu Cys Phe Glu Lys Val Asn Val Ala Gly Asp Thr Phe Gly Asn Cys 20 115 120 125 Gly Lys Asp 130
- 25 SEQ ID NO:15:

30

35

40

45

50

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1183 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

					CT G hr G												46
5																TGT	94
	Asn	Arg	Arg	Glu	Leu	Asp	Arg	Tyr	Leu	Gln	Ser	Gly	Gly	Gly	Met	Сув	
	CTC	TCC	AAC	ATG	CCA	GAC	ACC	AGG	ATG	TTG	TAT	GGA	GGC	CGG	AGG	TGT	142
10	Leu	Ser	Asn	Met	Pro	Asp	Thr	Arg	Met	Leu	Tyr	Gly	Gly	Arg	Arg	Cys	
,,,	GGG	AAC	GGG	TAT	CTG	GAA	GAT	GGG	GAA	GAG	TGT	GAC	TGT	GGA	GAA	GAA	190
	Gly	Asn	Gly	Tyr	Leu	Glu	Asp	Gly	Glu	Glu	Сув	Asp	Сув	Gly	Glu	Glu	
	GAG	GAA	TGT	AAC	AAC	ccc	TGC	TGC	AAT	GCC	TCT	AAT	TGT	ACC	CTG	AGG	238
15	Glu	Glu	Сув	Asn	Asn	Pro	Сув	Cys	Asn	Ala	Ser	Asn	Сув	Thr	Leu	Arg	
	CCG	GGG	GCG	GAG	TGT	GCT	CAC	GGC	TCC	TGC	TGC	CAC	CAG	TGT	AAG	CTG	286
	Pro	Gly	Ala	Glu	Сув	Ala	His	Gly	Ser	Сув	Cys	His	Gln	Cys	Lys	Leu	
20	TTG	GCT	CCT	GGG	ACC	CTG	TGC	CGC	GAG	CAG	GCC	AGG	CAG	TGT	GAC	CTC	334
					Thr												
	CCG	GAG	TTC	TGT	ACG	GGC	AAG	TCT	CCC	CAC	TGC	CCT	ACC	AAC	TTC	TAC	382
25					Thr												
25																	

	CAG	ATG	GAT	GGT	ACC	CCC	TGT	GAG	GGC	GGC	CAG	GCC	TAC	TGC	TAC	AAC	430
	Gln	Met	Asp	Gly	Thr	Pro	Сув	Glu	Gly	Gly	Gln	Ala	Tyr	Сув	Tyr	Asn	
5																CCC	478
	Gly	Met	Сув	Leu	Thr	Tyr	Gln	Glu	Gln	Сув	Gln	Gln	Leu	Trp	Gly	Pro	
					GCC												526
10	Gly	Ala	Arg	Pro	Ala	Pro	Asp	Leu	Сув	Phe	Glu	Lys	Val	Asn	Val	Ala	
	-	_		-	-											AGG	574
	_				_											Arg	
15																AGC	622
15	-	_			Arg	-											
					CCC												670
20				_	Pro												
20																TAC	718
					Gly												
					GAG												766
25	_	_			Glu												014
																TGC	814
				_	TTC	_										Сув	862
30					Phe												002
	Ī									_						GGC	910
					Cys												710
35	_		-		-											GAC	958
																Asp	
		•				-2-				•		•	•			•	•
					CCC												1006
40		_			Pro												
					TTG												1054
					Leu												
45					AAC												1102
45	-	_			Asn												
					CAA												1150
50		_			Gln							Arg	val	ser	GID	ASN	1103
50					CAT												1183
	Ser	Gly	Thr	Gly	His	Ala	Asn	Pro	Thr	Phe	Lys						

SEQ ID NO:16:

55

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 394 amino acids

- (B) TYPE: amino acid (D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

	Gly 1	Ala	Ala	Thr	Gly 5	His	Pro	Phe	Pro	Lys 10	Val	Phe	Asn	Gly	Cys 15	Asn
10	_			Leu 20	_				25				_	30		
			35	Pro	_			40		_	_	_	45		_	_
15		50		Leu			55					60				
	65			Asn		70					75					80
20	_			Cys	85					90					95	
				100 Thr		_			105					110		
			115	Thr	_	_		120					125			
25		130	_	Thr		_	135					140				
	145			Ala	_	150					155					160
30		_		Gly	165					170					175	
	Cys	Asn	Met	180 Arg	Asp	Ala	Lys	Сув	185 Gly	Lys	Ile	Gln	Cys	190 Gln	Ser	Ser
35			195 Arg	Pro	Leu	Glu		200 Asn	Ala	Val	Pro		205 Asp	Thr	Thr	Ile
33	Ile	210 Met	Asn	Gly	Arg		215 Ile	Gln	Сув	Arg		220 Thr	His	Val	Tyr	Arg 240
	Gly	Pro	Glu	Glu	Glu 245	230 Gly	Авр	Met	Leu	Asp 250	235 Pro	Gly	Leu	Val	Met 255	
40	Gly	Thr	Lys	Сув 260		Tyr	Asn	His	Ile 265		Leu	Glu	Gly	Gln 270		Arg
	Asn		275					280	_				285			
45		290					295					300				
	305			Phe		310					315					320
50	_			Pro	325					330					335	
-				Leu 340					345					350		
			355	Asn Gl'n				360					365			
55	_	370	_	His			375				7	380				~~~
	385	TILL	GLY	MAD	~14	390		****		2,0						

	SEQ ID NO:17:
	(i) SEQUENCE CHARACTERISTICS:
5	<ul><li>(A) LENGTH: 624 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: not relevant</li><li>(D) TOPOLOGY: linear</li></ul>
10	(ii) MOLECULE TYPE: cDNA (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vii) IMMEDIATE SOURCE:
15	(B) CLONE:CLONE TM
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:
20	
<i>2</i> 5	
30	•

	GC															CAA	47
		Thr	гув	Сув	Ala	ASP	GIÀ	гуя	116	Cys	Leu	ABII	Aty	GIII	Cys	Gln	
5																GGC	95
	Ast	ı Ile	: Sez	· Val	. Phe	Gly	Val	. His	g Glu	Cya	Ala	Met	: Gln	Cys	His	Gly	
	AGA	A GGG	GTC	TGC	: AAC	: AAC	AGG	AAG	AAC	TGC	CAC	TGC	: GAG	GCC	CAC	TGG	143
	Arg	g Gly	v Val	Cys	. Asn	Asn	Arg	Lye	aaA s	Cys	His	Cys	Glu	Ala	Hie	Trp	
10																	
																AGC	191
	Ala	a Pro	Pro	Phe	: Cys	Asp	гув	. PDE	e Gry	Pne	GIY	GIZ	, ser	The	AS	Ser	
																CTG	239
15	Gly	Pro	Ile	Arg	Gln	Ala	Asp	Asr	Gln	Gly	Leu	Thr	· Ile	Gly	Ile	Leu	
	G TY	ב אכנ	י איני	· care	: TCT	· (-11-1	. CT-1		י פכם	GGA	TT	· GTG	: СТТ	TAT '	CTC	AAA :	287
																Lys	
					-												
20																ATT	335
	Arg	Lys	Thr	Lev	Ile	Arg	Leu	Lev	ı Phe	Thr	Asn	Lys	Lys	Thr	Thi	: Ile	
	GAZ	AAA	CT	A AGG	TGI	GTG	CGC	: cc1	TCC	CGG	CCA	ccc	: CGI	GGC	TTC	CAA	383
																Gln	
25		_															
																AAG	431
	Pro	Cys	Glr	ı Ala	His	Leu	Gly	His	Leu	GLY	rya	GTA	r Leu	Met	Arg	l FAs	
	ccc	CCA	GA7	TCC	TAC	: CCA	CCG	AAG	GAC	AAT	ccc	: AGG	AGA	TTG	CTC	CAG	479
30																Gln	
00																	
																CCT	527
	Cys	3 Glr	Ası	ı Val	. Asp	TTE	ser	Arg	Pro	Leu	ASI	GTA	Leu	ASI	val	Pro	
or.																CCA	575
35	Glr	ı Pro	Glr	ı Sez	Thr	Gln	Arg	[Va]	Lev	Pro	Pro	Leu	His	Arg	Ala	Pro	
					-					· CTC	י ככי	CCC		י כריז	CCI	CTT	623
	CG:	ו הוא	Dec	COL	. Gic	DTC	Ala	Arc	r Dro	Lev	Pro	Ala	Lvs	Pro	Ala	Leu	023
	~~	, mr		, 561			, ,,,,,		,				, -				
40	A																624

## SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 207 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

50

45

- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

	Thr 1	Lys	Сув	Ala	Asp 5	Gly	Lys	Ile	Сув	Leu 10	Asn	Arg	Gln	Сув	Gln 15	Asn
5	Ile	Ser	Val	Phe 20	Gly	Val	His	Glu	Сув 25	Ala	Met	Gln	Сув	His 30	Gly	Arg
	Gly	Val	Сув 35	Asn	Asn	Arg	Lys	Asn 40	Сув	His	Сув	Glu	Ala 45	His	Trp	Ala
	Pro	Pro 50	Phe	Сув	qeA	ГÀа	Phe 55	Gly	Phe	Gly	Gly	Ser 60	Thr	qaA	Ser	Gly
10	Pro 65	Ile	Arg	Gln	Ala	Asp 70	Asn	Gln	Gly	Leu	Thr 75	Ile	Gly	Ile	Leu	Val 80
	Thr	Ile	Leu	Сув	Leu 85	Leu	Ala	Ala	Gly	Phe 90	Val	Val	Tyr	Leu	Lys 95	Arg
15	Lys	Thr	Leu	Ile 100	Arg	Leu	Leu	Phe	Thr 105	Asn	Lys	Lys	Thr	Thr 110	Ile	Glu
•	Lys	Leu	Arg 115	Cys	Val	Arg	Pro	Ser 120	Arg	Pro	Pro	Arg	Gly 125	Phe	Gln	Pro
	Суз	Gln 130	Ala	His	Leu	Gly	His 135	Leu	Gly	Lys	Gly	Leu 140	Met	Arg	Lys	Pro
20	Pro 145	Asp	Ser	Tyr	Pro	Pro 150	Lys	Asp	Asn	Pro	Arg 155	Arg	Leu	Leu	Gln	Cys 160
	Gln	Asn	Val	Asp	Ile 165	Ser	Arg	Pro	Leu	Asn 170	Gly	Leu	Asn	Val	Pro 175	Gln
25	Pro	Gln	Ser	Thr 180	Gln	Arg	Val	Leu	Pro 185	Pro	Leu	His	Arg	Ala 190	Pro	Arg
	Ala	Pro	Ser 195	Val	Pro	Ala	Arg	Pro 200	Leu	Pro	Ala	Lys	Pro 205	Ala	Leu	

#### SEQ ID NO:19:

30

35

40

45

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2669 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vii) IMMEDIATE SOURCE
  - (B) CLONE:JM109 (pMel  $\beta$ -24C) JM109 (pMel  $\beta$ -24N)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:
- 50 C GGA GCT GCC ACT GGG CAC CCC TTT CCC AAA GTG TTC AAT GGA TGC
  Gly Ala Ala Thr Gly His Pro Phe Pro Lys Val Phe Asn Gly Cys
  - AAC AGG AGG GAG CTG GAC AGG TAT CTG CAG TCA GGT GGT GGA ATG TGT 94

	Asn	Arg	Arg	Glu	Leu	Asp	Arg	Tyr	Leu	Gln	Ser	Gly	Gly	Gly	Met	Cys		
	C-TIC	TCC	AAC	STA	CCA	GAC	ACC	AGG	ATG	TTG	ТАТ	GGA	GGC	CGG	AGG	TGT		142
5	_															Сув		
	ccc	AAC	ccc	ТДТ	CTG	CAA	GAT	GGG	GAA	GAG	тут	GAC	тат	GGA	GAA	GAA		190
																Glu		
	<b>U</b> -1		,	-,-							•	•	•	•				
10																AGG	:	238
	Glu	Glu	Сув	Asn	Asn	Pro	Сув	CAa	Asn	Ala	Ser	Asn	Сув	Thr	Leu	Arg		
												<b>63.6</b>	~~~	mom.		como.		206
																CTG Leu	•	286
15	PIO	GTA	Ala	GIU	Cys	ALG	ure	GIY	361	cys	Cys		<b>G111</b>	Cyb	275	200		
	TTG	GCT	CCT	GGG	ACC	CTG	TGC	CGC	GAG	CAG	GCC	AGG	CAG	TGT	GAC	CTC	;	334
																Leu		
~~																TAC		382
20	Pro	GIu	Pne	сув	Thr	GIY	гув	ser	Pro	HIS	Сув	PIO	Int	ASII	Pne	TYL		
	CAG	ATG	GAT	GGT	ACC	CCC	TGT	GAG	GGC	GGC	CAG	GCC	TAC	TGC	TAC	AAC		430
																Asn		
			_	_														
25																CCC	•	478
	Gly	Met	Cys	Leu	Thr	Tyr	Gln	Glu	Gln	Сув	Gln	Gln	Leu	Trp	Gly	Pro		
			CC.		000	COR	CNC	CALCA	TCC	<b>₩</b>	CNC	220	GTG.	አልጥ	GTG	GCA	(	526
																Ala	•	<i></i>
30	UL,		· 5						-3-			-,-						
																AGG	!	574
	Gly	qaA	Thr	Phe	Gly	Asn	Cys	Gly	Lys	Asp	Met	Asn	Gly	Glu	His	Arg		
													~~~	mam	~~	7.00		
																AGC Ser	,	622
35	Lys	Cys	VDII	MEC	AL 9	~DP	ALG.	Lya	CyG	U.J	275			-,	· · · ·			
	TCT	GAG	GCC	CGG	CCC	CTG	GAG	TCC	AAC	GCG	GTG	CCC	ATT	GAC	ACC	ACT		670
									Asn									
																m. 0		
40	ATC	ATC	ATG	AAT	GGG	AGG	CAG	ATC	CAG	TGC	CGG	GGC	ACC The	CAC	Ual	TAC Tyr		718
	iie	116	mec	ABN	GIY	Arg	GIH	116	GIII	Cys	Æ	GLY	1111	1110	Val	- 7-		
	CGA	GGT	CCT	GAG	GAG	GAG	GGT	GAC	ATG	CTG	GAC	CCA	GGG	CTG	GTG	ATG		766
	Arg	Gly	Pro	Glu	Glu	Glu	Gly	Asp	Met	Leu	Asp	Pro	Gly	Leu	Val	Met		
45																		
	ACT	GGA	ACC	AAG	TGT	GGC	TAC	AAC	CAT	ATT	TGC	CTT	GAG	GGG	CAG	TGC	1	814
	Thr	GIA	Thr	Lys	сув	GIA	ıyr	ASI	His	TTE	Сув	rea	GIU	GIÅ	GIII	Cys		
	AGG	AAC	ACC	TCC	TTC	TTT	GAA	ACT	GAA	GGC	TGT	GGG	AAG	AAG	TGC	AAT	1	862
50									Glu									
50																		
									CAG								:	910
	Gly	His	Gly	Val	Сув	Asn	Asn	Asn	Gln	ASN	Сув	H18	CAa	ren	PTO	grA		
	TCC	ccc	cce	ccc	district.	TCC	ממ	ልሮል	CCG	GGC	CAC	GGG	GGC	AGT	ATC	GAC	. (958
55	Tro	Ala	Pro	Pro	Phe	Cys	Asn	Thr	Pro	Gly	His	Gly	Gly	Ser	Ile	Asp		
	- 4-					-				-		-	_					

															GGA Gly	GTG Val	1006
5	TTG	GTG	GCC	ATC	TTG	GTG	CTG	GCG	GTC	CTC	ATG	CTG	ATG	TAC	TAC	TGC	1054
	Leu	Val	Ala	Ile	Leu	Val	Leu	Ala	Val	Leu	Met	Leu	Met	Tyr	Tyr	Сув	
10														_	CTC Leu		1102
															CAG Gln		1150
15															CGG		1198
13		_		_							_				Arg		
						_									GGC		1246
20				CAC His										GCT	GACCI	ICA	1295
																GGACC	1355
25	GCC#	CCGC	CT (CTCC	LACTO	T GO	GAAC	AGCT	CAC	CACCO	CCA	aagi	GGTG	AC (CACC	TCCACG	1415 1475
	ACC	CTGC	AG 1	rgctc	CCCA	ig C	GCC1	rgccz	ACC	TTC	AGCG	TGTC	CACI	CT (TCC1	CCCGC CCTCA CCTTC	1535 1595 1655
30	TGC	GGGG	AT 1	TTGG#	CAGI	T TI	TCTC	:GCCC	GGG	GAAG	TCA	TCTA	CAAT	CAA (BACCO	ACCGA AGGGC	1715 1775
	CCTC	GCTC	TG A	ACAAT	GCCA	T CC	CTCI	rccgo	CAG	GTG	ATG	AGAC	CTGG	AC (CTG	CTGCC	1835 1895
35	GCCZ	ACGI	CA C	CTGC	GTGA	A CA	AGCA	CCTG	ccc	ATCA	AAG	TGTC	GGAC	ecc o	GAGCO	CTGTG CAGCCC CATTCC	1955 2015 2075
																GAGAG CCTCT	2135 2195
40	ATC	TCCI	CA C	TGTC	ACCA	T GG	TGCA	TGGG	AAG	GAGG	AGG	GCCI	GATO	CT (TTTG	ACCAA	2255 2315
	ACCA	TGCG	TG 1	rggac	ATTC	C TG	CCCI	GGGC	GTG	AGCG	TCA	CCTI	CAA1	GG (CAAG	CCACC TCTTC CCTGC	2375 2435 2495
45	ACCA	ACAA	ICC F	AGAGG	GACG	A CI	GTCI	CCAG	CGG	GACG	GAA	CCAC	TGCC	GC (CAGTI	GCAAG CGACT	2555 2615
73	GGCA	CACC	CC C	CACI	CCA	cc cc	CCGC	AGCC	: cc	GTGT	CTA	GCAC	ACCC	AC (CCCG		2669

SEQ ID NO:20:

50

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 427 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

	1			Thr	5					10					15	
5				Leu 20					25					30		
			35	Pro	_			40					45			
		50		Leu		_	55					60				
10	65	_		Asn		70					75					80
	_			Суз	85					90					95	
15			_	Thr 100					105					110		
			115	Thr	_			120					125			
20		130	_	Thr Thr			135					140				
	145	-			_	150					155					160
		_		Ala	165					170					175	
25	_			Gly 180					185					190		
	-		195	Arg Pro				200					205			
		210	_				215					220				
30	225			Gly	_	230					235					240
	_			Glu Cys	245	_				250					255	
35				260 Phe	_				265					270		
			275	Cys				280					285			
40		290					295					300				
40	305			Phe		310					315					320
	_			Pro	325					330					335	
45				Leu 340					345					350		
			355	Asn				360					365			
	Lys	Leu 370	Arg	Gln	Gln	Phe	Ser 375	Сув	Pro	Phe	Arg	Val 380	Ser	Gln	Asn	Ser
50	Gly 385	Thr	Gly	His	Ala	Asn 390	Pro	Thr	Phe	Lys	Pro 395	Glu	Phe	Arg	Ala	Pro 400
	His	Ser	Pro	His	His 405	His	Asp	Lys	Gly	His 410	Gln	Phe	His	Gly	His 415	Thr
55	Leu	Leu	His	Ser 420	Gly	Двр	qeA	Pro	Asp 425	Pro	His					

SEQ ID N0:21:

	(i) SEQUENCE CHARACTERISTICS:
5	(A) LENGTH: 1483 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: not relevant(D) TOPOLOGY: linear
10	(ii) MOLECULE TYPE: cDNA (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vii) IMMEDIATE SOURCE
	(B) CLONE:JM109 (pMel α-25C)
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:
20	
25	
30	
35	
40	
45	
50	

	GAT	GGG	CAC	TCA	TGT	CAG	GAT	GTG	GAC	GGC	TAC	TGC	TAC	AAT	GGC	ATC	48
	Asp	Gly	His	Ser	Cys	Gln	Asp	Val	Asp	Gly	Tyr	Сув	Tyr	Asn	Gly	Ile	
5																	
•									-							GAT	144
	Lys	Pro	Ala	Pro	GIA	Ile	Cys	Phe	GIu	Arg	Val	Asn	Ser	Ala	Gly	Asp	
	CCT	TAT	GGC	AAC	TGT	GGC	AAA	GTC	TCG	AAG	AGT	TCC	TTT	GCC	AAA	TGC	192
10	Pro	Tyr	Gly	Asn	Cys	Gly	Lys	Val	Ser	Lys	Ser	Ser	Phe	Ala	Lys	Сув	
	GAG	ATG	AGA	GAT	GCT	AAA	TGT	GGA	AAA	ATC	CAG	TGT	CAA	GGA	GGT	GCC	240
																Ala	
	D.C.C.	~~~	CCA	CTC.	እጥጥ	CCT	n.c.c	n n m	ccc	Calen	maa	200	C	101	220	ATC	200
15																Ile	288
	Ser	Arg	PIO	Val	116	GIY	Int	ABII	Ala	Val	Ser	rre	GIU	Thr	ABN	TTE	
						GGC											336
	Pro	Leu	Gln	Gln	Gly	Gly	Arg	Ile	Leu	Сув	Arg	Gly	Thr	His	Val	Tyr	
20	TTG	GGC	GAT	GAC	ATG	CCG	GAC	CCA	GGG	CTT	GTG	CTT	GCA	GGC	ACA	AAG	384
	Leu	Gly	Asp	Asp	Met	Pro	qaA	Pro	Gly	Leu	Val	Leu	Ala	Gly	Thr	Lys	
	TGT	GCA	GAT	GGA	AAA	ATC	TGC	CTG	AAT	CGT	CAA	TGT	CAA	ТАА	TTA	ACT	432
						Ile											
25																	
	-	_														GTG	480
	Val	Phe	Gly	Val	His	Glu	Cys	Ala	Met	Gln	Cys	His	Gly	Arg	Gly	Val	
	TGC	AAC	AAC	AGG	AAG	AAC	TGC	CAC	TGC	GAG	GCC	CAC	TGG	GCA	CCT	ccc	528
30	Сув	Asn	Asn	Arg	Lys	neA	Сув	His	Cys	Glu	Ala	His	Trp	Ala	Pro	Pro	
	ጉጥ ር	ጥርም	GAC	DAG	Total	GGC	طمامك	GGA	GGA	NGC	ACA	GAC	AGC	ccc	CCC	ATC	576
						Gly											370
		Cys	-Gp	275		<u>-</u>		0-3	OI,	001		νωp	JCI	017	- 10	110	
35	CGG	CAA	GCA	GAA	GCA	AGG	CAG	GAA	GCT	GCA	GAG	TCC	AAC	AGG	GAG	CGC	624
	Arg	Gln	Ala	Glu	Ala	Arg	Gln	Glu	Ala	Ala	Glu	Ser	Asn	Arg	Glu	Arg	
	GGC	CAG	GGC	CAG	GAG	ccc	GTG	GGA	TCG	CAG	GAG	САТ	GCG	тст	ACT	GCC	672
						Pro											0.2
40		~	~~}	~~	J. 4			 3		7-4	J. 4	-1	u				
	TCA	CTG	ACA	CTC	ATC	TGA	GCCC	TCCC	AT G	ACAT	GGAG	A CC	GTGA	CCAG	;		720
	Ser	Leu	Thr	Leu	Ile	*											

	TGCTGCTGCA	GAGGAGGTCA	CGCGTCCCCA	AGGCCTCCTG	TGACTGGCAG	CATTGACTCT	780
		CATCGTTTCC					840
5	_	CTACCAGGCA					900
	GTTGAGCTTC	TGCTAAAACA	TGGACATGCT	TCAGTGCTGC	TCCTGAGAGA	GTAGCAGGTT	960
	ACCACTCTGG	CAGGCCCCAG	CCCTGCAGCA	AGGAGGAAGA	GGACTCAAAA	GTCTGGCCTT	1020
	TCACTGAGCC	CCCACAGCAG	TGGGGGAGAA	GCAAGGGTTG	GGCCCAGTGT	CCCCTTTCCC	1080
	CAGTGACACC	TCAGCCTTGG	CAGCCCTGAT	GACTGGTCTC	TGGCTGCAAC	TTAATGCTCT	1140
10	GATATGGCTT	TTAGCATTTA	TTATATGAAA	ATAGCAGGGT	TTTAGTTTTT	AATTTATCAG	1200
	AGACCCTGCC	ACCCATTCCA	TCTCCATCCA	AGCAAACTGA	ATGGCATTGA	AACAAACTGG	1260
	AGAAGAAGGT	AGGAGAAAGG	GCGGTGAACT	CTGGCTCTTT	GCTGTGGACA	TGCGTGACCA	1320
	GCAGTACTCA	GGTTTGAGGG	TTTGCAGAAA	GCCAGGGAAC	CCACAGAGTC	ACCAACCCTT	1380
	CATTTAACAA	GTAAGAATGT	TAAAAAGTGA	AAACAATGTA	AGAGCCTAAC	TCCATCCCCC	1440
15	GTGGCCATTA	CTGCATAAAA	TAGAGTGCAT	CCCGCCCGAA	TTC		1483

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 230 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20

25

30

35

40

45

50

55

(ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Asp Gly His Ser Cys Gln Asp Val Asp Gly Tyr Cys Tyr Asn Gly Ile 10 Cys Gln Thr His Glu Gln Gln Cys Val Thr Leu Trp Gly Pro Gly Ala 20 5 25 Lys Pro Ala Pro Gly Ile Cys Phe Glu Arg Val Asn Ser Ala Gly Asp Pro Tyr Gly Asn Cys Gly Lys Val Ser Lys Ser Ser Phe Ala Lys Cys 10 Glu Met Arg Asp Ala Lys Cys Gly Lys Ile Gln Cys Gln Gly Gly Ala 70 75 Ser Arg Pro Val Ile Gly Thr Asn Ala Val Ser Ile Glu Thr Asn Ile 90 Pro Leu Gln Gln Gly Gly Arg Ile Leu Cys Arg Gly Thr His Val Tyr 15 100 105 Leu Gly Asp Asp Met Pro Asp Pro Gly Leu Val Leu Ala Gly Thr Lys 120 Cys Ala Asp Gly Lys Ile Cys Leu Asn Arg Gln Cys Gln Asn Ile Ser 135 140 Val Phe Gly Val His Glu Cys Ala Met Gln Cys His Gly Arg Gly Val 20 155 150 Cys Asn Asn Arg Lys Asn Cys His Cys Glu Ala His Trp Ala Pro Pro 165 170 Phe Cys Asp Lys Phe Gly Phe Gly Gly Ser Thr Asp Ser Gly Pro Ile 25 180 185 190 Arg Gln Ala Glu Ala Arg Gln Glu Ala Ala Glu Ser Asn Arg Glu Arg 200 Gly Gln Gly Gln Glu Pro Val Gly Ser Gln Glu His Ala Ser Thr Ala 215 220 Ser Leu Thr Leu Ile * 30 225 230

SEQ ID NO:23:

*3*5

40

45

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1569 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vii) IMMEDIATE SOURCE
 - (B) CLONE:JM109 (pMel α -26N)
- 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

	GGG	GAC	CTC	TGG	ATC	CCA	GTG	AAG	AGC	TTC	GAC	TCC	AAG	AAT	CAT	CCA	48
	Gly	Asp	Leu	Trp	Ile	Pro	Val	Lys	Ser	Phe	Asp	Ser	Lys	Asn	His	Pro	
5	GAA	GTG	CAC	AAT	ATT	CGA	CTA	CAA	CGG	GAA	AGC	AAA	GAA	CTG	ATC	ATA	96
																Ile	
	አልጥ	CTG	GAA	AGA	דממ	GAA	CCT	CTC	ATT	GCC	AGC	AGT	TTC	ACG	GAA	ACC	144
									Ile								
10	CAC	тат	CTG	CAA	GAC	GGT	ACT	GAT	GTC	TCC	CTC	GCT	CGA	AAT	TAC	ACG	192
																Thr	
	GGT	CAC	TGT	TAC	TAC	CAT	GGA	CAT	GTA	CGG	GGA	TAT	TCT	GAT	TCA	GCA	240
15																Ala	
	GTC	AGT	CTC	AGC	ACG	TGT	TCT	GGT	CTC	AGG	GGA	CTT	ATT	GGG	TTT	GAA	288
	Val	Ser	Leu	Ser	Thr	Сув	Ser	Gly	Leu	Arg	Gly	Leu	Ile	Gly	Phe	Glu	
20									ATG								336
	Asn	Glu	Ser	Tyr	Val	Leu	Glu	Pro	Met	Lys	Ser	Ala	Thr	Asn	Arg	Tyr	
																GGA	384
25	Lys	Leu	Phe	Pro	Ala	Lys	Lys	Leu	Lys	Ser	Val	Arg	Gly	Ser	Сув	Gly	
																CCA	432
	Ser	His	His	Asn	Thr	Pro	Asn	Leu	Ala	Ala	Lys	Asn	Val	Phe	Pro	Pro	
																GCA	480
30	Pro	Ser	Gln	Thr	Trp	Ala	Arg	Arg	His	Lys	Arg	Glu	Thr	Leu	Lys	Ala	
																CAG	528
	Thr	Lys	Tyr	Val	Glu	Leu	Val	Ile	Val	Ala	Asp	Asn	Arg	Glu	Phe	Gln	
35																ATT	576
	Arg	Gln	Gly	Lys	Asp	Leu	Glu	Lys	Val	Lys	Gln	Arg	Leu	Ile	GLu	Ile	
																GTG	624
40	Ala	Asn	His	Val	Asp	Lys	Phe	Tyr	Arg	Pro	Leu	Asn	Ile	Arg	He	Val	
	TTG	GTA							GAC								672
	•		 3		~ 1 · ·	77-7	-	2	2	Mak		1	~-~	Cow	37-7	C0~	

	CAG	GAC	CCA	TTC	ACC	AGC	CTC	CAT	GAA	TTT	CTG	GAC	TGG	AGG	AAG	ATG	720
	Gln	qaA	Pro	Phe	Thr	Ser	Leu	His	Glu	Phe	Leu	Asp	Trp	Arg	Lys	Met	
5																GGG	768
	•															Gly	
						ACC											816
10		_			_	Thr											
						TCT											864
	•					Ser											
15						GTG											912
.0						Val											
						ACA											960
	GIÀ	Met	Asn	HIS	ASP	Thr	Leu	Asp	Arg	GIA	Сув	Ser	Сув	GIN	met	Ala	
20	GTT	GAG	AAA	GGA	GGC	TGC	ATC	ATG	AAC	GCT	TCC	ACC	GGG	TAC	CCA	TTT	1008
			_	-	_	Сув											
																CTG	1056
25						Ser											
						GTG											1104
	Glù	Lys	Gly	Met	Gly	Val	Суз	Leu	Phe	Asn	Leu	Pro	Glu	Val	Arg	G1 <i>n</i>	
	TCT	TTC	GGG	GGC	CAG	AAG	TGT	GGG	AAC	AGA	TTT	GTG	GAA	GAA	GGA	GAG	1152
30						Lys											
	a. a		~~~	mam	~~~	GAG	CCR	030	CAR	m/cm	D-1707	227	cec	TYC	TYCC	224	1200
						Glu											1200
		_	_	_													
35						CTG											1248
	Ala	Thr	Thr	Cys	Thr	Leu	Lys	Pro	Asp	Ala	Val	СУВ	Ala	Hls	GIY	Leu	
	TGC	TGT	GAA	GAC	TGC	CAG	CTG	AAG	CCT	GCA	GGA	ACA	GCG	TGC	AGG	GAC	1296
40	Суз	Сув	Glu	Asp	Сув	Gln	Leu	Lys	Pro	Ala	Gly	Thr	Ala	Сув	Arg	Asp	
	TCC	AGC	AAC	TCC	TGT	GAC	CTC	CCA	GAG	TTC	TGC	ACA	GGG	GCC	AGC	CCT	1344
	Ser	Ser	Asn	Ser	Cys	qeA	Leu	Pro	Glu	Phe	CAa	Thr	Gly	Ala	Ser	Pro	
	CAC	TGC	CCA	GCC	AAC	GTG	TAC	CTG	CAC	GAT	GGG	CAC	TCA	TGT	CAG	GAT	1392
45						Val											
													a.a	~~~	~~	CNC	1440
						TAC Tyr											1440
	AGT	чер	GIA	TÄL	cys	TAT	WO!!	GIA	115	C10	GIII	****	***				
50						GGA											1488
	Сув	Val	Thr	Leu	Trp	Gly	Pro	Gly	Ala	Lys	Pro	Ala	Pro	Gly	Ile	Сув	
	TTT	GAG	AGA	GTC	AAT	TCT	GCA	GGT	GAT	CCT	TAT	GGC	AAC	TGT	GGC	AAA	1536
						Ser											
55			_					-	_			-					

1569

GTC TCG AAG AGT TCC TTT GCC AAA TGC GAG ATG

Val Ser Lys Ser Ser Phe Ala Lys Cys Glu Met

	SEQ ID NO:24:
	(i) SEQUENCE CHARACTERISTICS:
10	(A) LENGTH: 523 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear
15	(ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:
20	
25	
30	
35	
40	
4 5	
50	
55	

Gly Asp Leu Trp Ile Pro Val Lys Ser Phe Asp Ser Lys Asn His Pro Glu Val Leu Asn Ile Arg Leu Gln Arg Glu Ser Lys Glu Leu Ile Ile Asn Leu Glu Arg Asn Glu Gly Leu Ile Ala Ser Ser Phe Thr Glu Thr His Tyr Leu Gln Asp Gly Thr Asp Val Ser Leu Ala Arg Asn Tyr Thr Gly His Cys Tyr Tyr His Gly His Val Arg Gly Tyr Ser Asp Ser Ala Val Ser Leu Ser Thr Cys Ser Gly Leu Arg Gly Leu Ile Gly Phe Glu Asn Glu Ser Tyr Val Leu Glu Pro Met Lys Ser Ala Thr Asn Arg Tyr Lys Leu Phe Pro Ala Lys Lys Leu Lys Ser Val Arg Gly Ser Cys Gly Ser His His Asn Thr Pro Asn Leu Ala Ala Lys Asn Val Phe Pro Pro Pro Ser Gln Thr Trp Ala Arg Arg His Lys Arg Glu Thr Leu Lys Ala Thr Lys Tyr Val Glu Leu Val Ile Val Ala Asp Asn Arg Glu Phe Gln Arg Gln Gly Lys Asp Leu Glu Lys Val Lys Gln Arg Leu Ile Glu Ile Ala Asn His Val Asp Lys Phe Tyr Arg Pro Leu Asn Ile Arg Ile Val Leu Val Gly Val Glu Val Trp Asn Asp Met Asp Lys Cys Ser Val Ser Gln Asp Pro Phe Thr Ser Leu His Glu Phe Leu Asp Trp Arg Lys Met Lys Leu Leu Pro Arg Lys Ser His Asp Asn Ala Gln Leu Val Ser Gly Val Tyr Phe Gln Gly Thr Thr Ile Gly Met Ala Pro Ile Met Ser Met Cys Thr Ala Asp Gln Ser Gly Gly Ile Val Met Asp His Ser Asp Asn Pro Leu Gly Ala Ala Val Thr Leu Ala His Glu Leu Gly His Asn Phe Gly Met Asn His Asp Thr Leu Asp Arg Gly Cys Ser Cys Gln Met Ala Val Glu Lys Gly Gly Cys Ile Met Asn Ala Ser Thr Gly Tyr Pro Phe

45

5

10

15

20

25

30

35

40

50

Pro Met Val Phe Ser Ser Cys Ser Arg Lys Asp Leu Glu Thr Ser Leu Glu Lys Gly Met Gly Val Cys Leu Phe Asn Leu Pro Glu Val Arg Glu Ser Phe Gly Gly Gln Lys Cys Gly Asn Arg Phe Val Glu Glu Gly Glu Glu Cys Asp Cys Gly Glu Pro Glu Glu Cys Met Asn Arg Cys Cys Asn Ala Thr Thr Cys Thr Leu Lys Pro Asp Ala Val Cys Ala His Gly Leu Cys Cys Glu Asp Cys Gln Leu Lys Pro Ala Gly Thr Ala Cys Arg Asp Ser Ser Asn Ser Cys Asp Leu Pro Glu Phe Cys Thr Gly Ala Ser Pro His Cys Pro Ala Asn Val Tyr Leu His Asp Gly His Ser Cys Gln Asp Val Asp Gly Tyr Cys Tyr Asn Gly Ile Cys Gln Thr His Glu Gln Gln Cys Val Thr Leu Trp Gly Pro Gly Ala Lys Pro Ala Pro Gly Ile Cys Phe Glu Arg Val Asn Ser Ala Gly Asp Pro Tyr Gly Asn Cys Gly Lys Val Ser Lys Ser Ser Phe Ala Lys Cys Glu Met

SEQ ID NO:25:

5

10

15

20

25

30

35

40

45

50

- (i) SEQUÊNCE CHARACTERISTICS:
 - (A) LENGTH: 2404 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA
 - (iii) HYPOTHETICAL: NO
 - (iv) ANTI-SENSE: NO
 - (vii) IMMEDIATE SOURCE
 - (B) CLONE:JM109 (pMel β-24C)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

	TGC	TGC	CAC	CAG	TGT	AAG	CTG	TTG	GCT	CCT	GGG	ACC	CTG	TGC	CGC	GAG	48
	Сув	Сув	His	Gln	Суз	Lys	Leu	Leu	Ala	Pro	Gly	Thr	Leu	Сув	Arg	Glu	
5	CAG	GCC	AGG	CAG	TGT	GAC	CTC	CCG	GAG	TIC	TGT	ACG	GGC	AAG	TCT	CCC	96
	Gln	Ala	Arg	Gln	Сув	Asp	Leu	Pro	Glu	Phe	Сув	Thr	Gly	Lys	Ser	Pro	
	CAC	TGC	CCT	ACC	AAC	TTC	TAC	CAG	ATG	GAT	GGT	ACC	CCC	TGT	GAG	GGC	144
	His	Сув	Pro	Thr	Asn	Phe	Tyr	Gln	Met	qaA	Gly	Thr	Pro	Cys	Glu	Gly	
10	GGC	CAG	GCC	TAC	TGC	TAC	AAC	GGC	ATG	TGC	СТС	ACC	TAC	CAG	GAG	CAG	192
				Tyr													
	TGC	CAG	CAG	CTG	TGG	GGA	ccc	GGA	GCC	CGA	CCT	GCC	ССТ	GAC	CTC	TGC	240
15				Leu													
	TTC	GAG	AAG	GTG	AAT	GTG	GCA	GGA	GAC	ACC	TTT	GGA	AAC	TGT	GGA	AAG	288

5

	Phe	Glu	Lys	Val	Asn	Val	Ala	Gly	Авр	Thr	Phe	Gly	Asn	Сув	Gly	Lys	
	GAC	ATG	ТАА	CCT	GAA	CAC	AGG	AAG	TGC	AAC	ATYZ	ACA	GAT	GCG	AAG	TGT	226
5																Сув	336
																AAC Asn	384
	GLY	цз	116	GIII	cys	GIII	261	Jer	GIU	VIG	ALG	PIO	Den	GIU	ser	ASII	
10	GCG	GTG	ccc	ATT	GAC	ACC	ACT	ATC	ATC	ATG	AAT	GGG	AGG	CAG	ATC	CAG	432
	Ala	Val	Pro	Ile	qeA	Thr	Thr	Ile	Ile	Met	Asn	Gly	Arg	Gln	Ile	Gln	
	TYCE	ccc	GGC	NCC.	CNC	GTC	TAC	CGA	CCT	CCT	CNG	020	CNC	com	CNG	ATG	400
													Glu				480
15	•						•	5	3					,			
													GGC				528
	Leu	qaA	Pro	Gly	Leu	Val	Met	Thr	Gly	Thr	ГÀЗ	Суз	Gly	Tyr	Asn	His	
	ATT	TGC	CTT	GAG	GGG	CAG	TGC	AGG	AAC	ACC	TCC	TTC	TTT	GAA	ACT	GAA	576
20																Glu	0.0
													AAC			CAG	624
	Q.J	Cys	Q.Y	Dy 3	273	-ys	7.514	O _L y	*****	Gry	VGI	cys	ASII	ABII	וומה	GIII	
25													TGC				672
	Asn	Cys	His	Cys	Leu	Pro	Gly	Trp	Ala	Pro	Pro	Phe	Сув	Asn	Thr	Pro	
	GGC	CAC	GGG	GGC	AGT	ATC	GAC	AGT	GGG	CCT	ATG	CCC	CCT	DAD	AGT	GTG	720
													Pro				,20
30							_										
													GTG Val				768
	GIY	PIG	Val	Vai	AIG	GIY	Val	reu	val	MIG	TIE	Den	val	rea	ALA	val	
	CTC	ATG	CTG	ATG	TAC	TAC	TGC	TGC	AGA	CAG	AAC	AAC	AAA	CTA	GGC	CAA	816
35	Leu	Met	Leu	Met	Tyr	Tyr	Сув	Сув	Arg	Gln	neA	Asn	Lys	Leu	Gly	Gln	
	כייור	DAG	CCC	TCA	сст	CTC	CCT	TCC	DAG	CTG	AGG	CAA	CAG	ጥጥር	ACT	TOT	864
	Leu																004
40													GCC				912
	PIO	Phe	Atg	Val	ser	GIII	MBII	Ser	GIY	Inc	GIY	nis	Ala	ASII	Pro	The	
	TTC	AAG	CCG	GAA	TTC	CGG	GCC	CCC	CAC	AGC	CCA	CAC	CAC	CAT	GAC	AAG	960
	Phe	Lys	Pro	Glu	Phe	Arg	Ala	Pro	His	Ser	Pro	His	His	His	Asp	Lys	
45	GGC	CAC	CAA	TTC	ראר	GGC	CAC	ACC	CTC	CTC	CAC	тст	GGG	GAC	GAC	cca	1008
													Gly				1008
	_												-	•	_		
		-		TGA	GCTG	ACCA	CA A	CAGC	CACT	A CA	ACTG	CAGC	CAC	TGGA	TCC		1060
50	Asp	Pro	HIS	•													
	ACGG	CCAC	CC T	GTCC	TCCA	c cc	CAGG	GACC	ACC	TGGA	TCC	TCAC	'AGAG	CC G	AGCA	CTATA	1120
															-	CAGCT	1180
																CGGTT	1240
55																TGCCA GCTTC	1300 1360
	wee1		I	J. C.C	1 G		1		GIC	CACH		CCL	NUND		WC 10	OC11C	7200

	CCCAGCTCCC	ACTTCTCTAC	TCCCTGCTTC	TGCAGGGCAT	TTGGACAGTT	TTTCTCGCCC	1420
	GGGGAAGTCA	TCTACAATAA	GACCGACCGA	GCCGGCTGCC	ATTTCTACGC	AGTGTGCAAT	1480
	CAGCACTGTG	ACATTGACCG	CTTCCAGGGC	GCCTGTCCCA	CCTCCCCACC	GCCAGTGTCC	1540
5	TCCGCCCCGC	TGTCCTCGCC	CTCCCCTGCC	CCTGGCTGTG	ACAATGCCAT	CCCTCTCCGG	1600
	CAGGTGAATG	AGACCTGGAC	CCTGGAGAAC	TGCACGGTGG	CCAGGTGCGT	GGGTGACAAC	1660
	CGTGTCGTCC	TGCTGGACCC	AAAGCCTGTG	GCCAACGTCA	CCTGCGTGAA	CAAGCACCTG	1720
	CCCATCAAAG	TGTCGGACCC	GAGCCAGCCC	TGTGACTTCC	ACTATGAGTG	CGAGTGCATC	1780
	TGCAGCATGT	GGGGCGGCTC	CCACTATTCC	ACCTTTGACG	GCACCTCTTA	CACCTTCCGG	1840
10	GGCAACTGCA	CCTATGTCCT	CATGAGAGAG	ATCCATGCAC	GCTTTGGGAA	TCTCAGCCTC	1900
	TACCTGGACA	ACCACTACTG	CACGGCCTCT	GCCACTGCCG	CTGCCGCCCG	CTGCCCCCGC	1960
	GCCCTCAGCA	TCCACTACAA	GTCCATGGAT	ATCGTCCTCA	CTGTCACCAT	GGTGCATGGG	2020
	AAGGAGGAGG	GCCTGATCCT	GTTTGACCAA	ATTCCGGTGA	GCAGCGGTTT	CAGCAAGAAC	2080
	GGCGTGCTTG	TGTCTGTGCT	GGGGACCACC	ACCATGCGTG	TGGACATTCC	TGCCCTGGGC	2140
15	GTGAGCGTCA	CCTTCAATGG	CCAAGTCTTC	CAGGCCCGGC	TGCCCTACAG	CCTCTTCCAC	2200
	AACAACACCG	AGGGCCAGTG	CGGCACCTGC	ACCAACAACC	AGAGGGACGA	CTGTCTCCAG	2260
	CGGGACGGAA	CCACTGCCGC	CAGTTGCAAG	GACATGGCCA	AGACGTGGCT	GGTCCCCGAC	2320
	AGCAGAAAGG	ATGGCTGCTG	GGCCCCGACT	GGCACACCCC	CCACTGCCAG	CCCCGCAGCC	2380
	CCGGTGTCTA	GCACACCCAC	CCCG				2404
20							

SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 339 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Cys Cys His Gln Cys Lys Leu Leu Ala Pro Gly Thr Leu Cys Arg Glu 10 Gln Ala Arg Gln Cys Asp Leu Pro Glu Phe Cys Thr Gly Lys Ser Pro 25 5 His Cys Pro Thr Asn Phe Tyr Gln Met Asp Gly Thr Pro Cys Glu Gly 40 Gly Gln Ala Tyr Cys Tyr Asn Gly Met Cys Leu Thr Tyr Gln Glu Gln 10 Cys Gln Gln Leu Trp Gly Pro Gly Ala Arg Pro Ala Pro Asp Leu Cys 75 Phe Glu Lys Val Asn Val Ala Gly Asp Thr Phe Gly Asn Cys Gly Lys 85 90 Asp Met Asn Gly Glu His Arg Lys Cys Asn Met Arg Asp Ala Lys Cys 15 105 Gly Lys Ile Gln Cys Gln Ser Ser Glu Ala Arg Pro Leu Glu Ser Asn 120 Ala Val Pro Ile Asp Thr Thr Ile Ile Met Asn Gly Arg Gln Ile Gln 135 140 Cys Arg Gly Thr His Val Tyr Arg Gly Pro Glu Glu Glu Gly Asp Met 20 150 155 Leu Asp Pro Gly Leu Val Met Thr Gly Thr Lys Cys Gly Tyr Asn His 170 165 Ile Cys Leu Glu Gly Gln Cys Arg Asn Thr Ser Phe Phe Glu Thr Glu 25 185 190 Gly Cys Gly Lys Lys Cys Asn Gly His Gly Val Cys Asn Asn Asn Gln 195 30 Asn Cys His Cys Leu Pro Gly Trp Ala Pro Pro Phe Cys Asn Thr Pro 215 Gly His Gly Gly Ser Ile Asp Ser Gly Pro Met Pro Pro Glu Ser Val 230 235 35 Gly Pro Val Val Ala Gly Val Leu Val Ala Ile Leu Val Leu Ala Val 245 250 Leu Met Leu Met Tyr Tyr Cys Cys Arg Gln Asn Asn Lys Leu Gly Gln 265 Leu Lys Pro Ser Ala Leu Pro Ser Lys Leu Arg Gln Gln Phe Ser Cys 40 280 Pro Phe Arg Val Ser Gln Asn Ser Gly Thr Gly His Ala Asn Pro Thr 295 300 Phe Lys Pro Glu Phe Arg Ala Pro His Ser Pro His His Asp Lys 310 315 45 Gly His Gln Phe His Gly His Thr Leu Leu His Ser Gly Asp Asp Pro 330 Asp Pro His 339 50

SEQ ID NO:27:

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 453 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO

5		(vii) li	MMED	IATE :	SOUR	CE								
3		(B) CLO	ONE :J	M109	(pMei	β-24Ν)						
		(xi) S	EQUE	NCE [DESC	RIPTIC	N: SE	Q ID I	NO:27	:				
10		GA GO										 	 	46
15		AGG Arg										 	 	94
													TGT Cys	142
20												-	GAA Glu	190
		GAA Glu					_				 	 	 	238
25		GGG Gly											CTG Leu	286
30		GCT Ala											 CTC Leu	334
35									_				TAC Tyr	382
													AAC Asn	430
40		ATG Met	-	-				GA						453
45	INF	ORM	ATION	FOR	SEQ II	D NO:	28:							
40		(i) SE	QUEN	ICE C	HARA	CTER	ISTIC	3 :						
50		(A) LEN B) TYF D) TO	PE: am	ino ac	id	acids							
			OLEC					(Q ID I	NO:28	:				

Gly Ala Ala Thr Gly His Pro Phe Pro Lys Val Phe Asn Gly Cys Asn Arg Arg Glu Leu Asp Arg Tyr Leu Gln Ser Gly Gly Met Cys Leu 25 20 Ser Asn Met Pro Asp Thr Arg Met Leu Tyr Gly Gly Arg Arg Cys Gly 40 Asn Gly Tyr Leu Glu Asp Gly Glu Glu Cys Asp Cys Gly Glu Glu Glu 60 Glu Cys Asn Asn Pro Cys Cys Asn Ala Ser Asn Cys Thr Leu Arg Pro 70 75 Gly Ala Glu Cys Ala His Gly Ser Cys Cys His Gln Cys Lys Leu Leu 90 Ala Pro Gly Thr Leu Cys Arg Glu Gln Ala Arg Gln Cys Asp Leu Pro 105 110 Glu Phe Cys Thr Gly Lys Ser Pro His Cys Pro Thr Asn Phe Tyr Gln Met Asp Gly Thr Pro Cys Glu Gly Gly Gln Ala Tyr Cys Tyr Asn Gly Met Cys Leu Thr Tyr Gln 150 145

25 Claims

5

10

15

20

- A soluble meltrin polypeptide which does not comprise a transmembrane domain or an intracellular domain and which comprises the amino acid sequence of Glu (No. 156) to Ile (No. 686) from the N-terminal in Fig. 15a - Fig. 15f.
- A polypeptide of claim 1 which comprises the amino acid sequence of Gly (No. 1) to Ile (No. 686) from the N-terminal in Fig. 15a Fig. 15f.
 - 3. A DNA comprising a base sequence encoding a polypeptide of claim 1 or 2.
- 4. A DNA of claim 3 which comprises the base sequence of No. 1 to No. 2058 from the 5' terminal in Fig. 15a Fig. 15f.
 - 5. A DNA of claim 3 which comprises the base sequence of No. 1 to No. 2848 from the 5' terminal in Fig. 15a Fig. 15f.
- 6. An antisense oligonucleotide which hybridizes with a part of the sequence of No. 1957 to No. 2848 from the 5' terminal in Fig. 15a Fig. 15f.
 - 7. An antisense oligonucleotide of claim 6 which inhibits the expression of the polypeptide of claim 1 or 2.
- 8. An antibody which recognizes the C-terminal region of a soluble meltrin wherein the C-terminal region is from amino acid No. 653 to No. 686 from the amino terminal in Fig. 15a Fig. 15f.
 - 9. An antibody of claim 8 which is a polyclonal antibody obtained from a mouse.
 - 10. 4. An antibody of claim 8 which is a monoclonal antibody.
 - 11. An antibody of claim 10 which is a monoclonal antibody obtained from a hybridoma using mouse spleen cells and lymphocytes.
 - 12. A method for the preparation of an antibody which method comprises:
 - immunizing an animal with a polypeptide of claim 1 or 2; and
 - obtaining an antibody from the immunized animal.

55

- 13. A vector comprising a DNA of any one of claims 3 to 5.
- 14. A transformant by the vector of claim 13.
- 5 15. A process for producing a polypeptide of claim 1 or 2, which process comprises culturing the appropriate transformant of claim 14.
 - 16. A medical composition comprising a polypeptide of claim 1 or 2, an antisense oligonucleotide of claim 6 or 7 or an antibody of any one of claims 8 to 11.
 - 17. Use of a polypeptide of claim 1 or 2, an antisense oligonucleotide of claim 6 or 7 or an antibody of any one of claims 8 to 11 for the manufacture of a medicament for treatment of a condition associated with unhealthy enhanced bone resorption.
- 15 18. Use according to claim 17, wherein the condition associated with unhealthy enhanced bone resorption is osteoporosis or hypercalsemia.
 - 19. Use of a polypeptide of claim 1 or 2, an antisense oligonucleotide of claim 6 or 7 or an antibody of any one of claims 8 to 11 for the manufacture of a medicament for preventing metastasis of cancer cells.

Patentansprüche

10

20

35

- Lösliches Meltrin-Polypeptid, nicht umfassend eine Transmembrandomäne oder eine intrazelluläre Domäne und umfassend die Aminosäuresequenz von Glu (Nr. 156) bis Ile (Nr. 686) von dem N-Terminus in Fig. 15a - Fig. 15f.
 - 2. Polypeptid nach Anspruch 1, umfassend die Aminosäuresequenz von Gly (Nr. 1) bis IIe (Nr. 686) von dem N-Terminus in Fig. 15a Fig. 15f.
- 30 3. DNA, umfassend eine Basensequenz, welche ein Polypeptid nach Anspruch 1 oder 2 kodiert.
 - 4. DNA nach Anspruch 3, umfassend die Basensequenz von Nr. 1 bis Nr. 2058 von dem 5'-Terminus in Fig. 15a Fig. 15f.
 - 5. DNA nach Anspruch 3, umfassend die Basensequenz von Nr. 1 bis Nr. 2848 von dem 5'-Terminus in Fig. 15a Fig. 15f.
 - 6. Antisense-Oligonucleotid, welches mit einem Teil der Sequenz von Nr. 1957 bis Nr. 2848 von dem 5'-Terminus in Fig. 15a Fig. 15f hybridisiert.
- Antisense-Oligonucleotid nach Anspruch 6, welches die Expression des Polypeptids nach Anspruch 1 oder 2 inhibiert.
 - 8. Antikörper, welcher die C-terminale Region eines löslichen Meltrins erkennt, wobei sich die C-terminale Region von Aminosäure Nr. 653 bis Nr. 686 von dem Aminoterminus in Fig. 15a Fig. 15f erstreckt.
- 45 9. Antikörper nach Anspruch 8, welcher ein polyklonaler Antikörper ist, der von einer Maus erhalten wurde.
 - 10. Antikörper nach Anspruch 8, welcher ein monoklonaler Antikörper ist.
- 11. Antikörper nach Anspruch 10, welcher ein monoklonaler Antikörper ist, erhalten aus einem Hybridom unter Verwendung von Mäusemilzzellen und Lymphozyten.
 - 12. Verfahren zur Herstellung eines Antikörpers, wobei das Verfahren umfasst:
 - Immunisieren eines Tieres mit einem Polypeptid nach Anspruch 1 oder 2; und
 - Erhalten eines Antikörpers aus dem immunisierten Tier.
 - 13. Vektor, umfassend eine DNA nach einem der Ansprüche 3 bis 5.

- 14. Transformant durch den Vektor nach Anspruch 13.
- 15. Verfahren zur Herstellung eines Polypeptids nach Anspruch 1 oder 2, wobei das Verfahren Kultivieren des geeigneten Transformanten nach Anspruch 14 umfasst.
- 16. Arzneimittel, umfassend ein Polypeptid nach Anspruch 1 oder 2, ein Antisense-Oligonucleotid nach Anspruch 6 oder 7 oder einen Antikörper nach einem der Ansprüche 8 bis 11.
- 17. Verwendung eines Polypeptids nach Anspruch 1 oder 2, eines Antisense-Oligonucleotids nach Anspruch 6 oder 7 oder eines Antikörpers nach einem der Ansprüche 8 bis 11 zur Herstellung eines Medikaments zur Behandlung eines Zustandes, der mit krankhafter gesteigerter Knochenresorption in Zusammenhang steht.
 - 18. Verwendung gemäß Anspruch 17, wobei der Zustand, der mit krankhafter gesteigerter Knochenresorption in Zusammenhang steht, Osteoporose oder Hyperkalzämie ist.
 - 19. Verwendung eines Polypeptids nach Anspruch 1 oder 2, eines Antisense-Oligonucleotids nach Anspruch 6 oder 7 oder eines Antikörpers nach einem der Ansprüche 8 bis 11 zur Herstellung eines Medikaments zur Verhinderung einer Metastasierung von Krebszellen.

Revendications

- Polypeptide de meltrine soluble, qui ne comporte pas un domaine transmembranaire ou un domaine intracellulaire et qui comporte la séquence d'acides aminés allant du résidu Glu n° 156 au résidu lle n° 686 du domaine N-terminal présenté sur les figures 15a à 15f.
- Polypeptide conforme à la revendication 1, qui comporte la séquence d'acides aminés allant du résidu Gly n° 1 au résidu lle n° 686 du domaine N-terminal présenté sur les figures 15a à 15 f.
- 30 3. ADN comportant une séquence de bases qui code un polypeptide conforme à la revendication 1 ou 2.
 - 4. ADN conforme à la revendication 3, qui comporte la séquence de bases allant de la base n° 1 à la base n° 2058 du domaine 5'-terminal présenté sur les figures 15a à 15f.
- 35 5. ADN conforme à la revendication 3, qui comporte la séquence de bases allant de la base n° 1 à la base n° 2848 du domaine 5'-terminal présenté sur les figures 15a à 15f.
 - 6. Oligonucléotide anti-sens qui s'hybride avec une partie de la séquence allant de la base n° 1957 à la base n° 2848 du domaine 5'-terminal présenté sur les figures 15a à 15f.
 - Oligonucléotide anti-sens conforme à la revendication 6, qui inhibe l'expression d'un polypeptide conforme à la revendication 1 ou 2.
- Anticorps reconnaissant un domaine C-terminal d'une meltrine soluble, lequel domaine C-terminal va du résidu
 d'acide aminé n° 653 3 au résidu n° 686 du domaine amino-terminal présenté sur les figures 15a à 15f.
 - 9. Anticorps conforme à la revendication 8, qui est un anticorps polyclonal obtenu chez une souris.
 - 10. Anticorps conforme à la revendication 8, qui est un anticorps monoclonal.
 - 11. Anticorps conforme à la revendication 10, qui est un anticorps monoclonal obtenu à partir d'un hybridome formé avec des cellules spléniques de souris et des lymphocytes.
 - 12. Procédé de préparation d'un anticorps, lequel procédé comporte :
 - le fait d'immuniser un animal avec un polypeptide conforme à la revendication 1 ou 2,
 - et le fait de récupérer un anticorps chez cet animal immunisé.

25

15

5

40

55

- 13. Vecteur comprenant un ADN conforme à l'une des revendications 3 à 5.
- 14. Organisme transformé avec un vecteur conforme à la revendication 13.

10

20

25

30

35

40

45

50

55

- 5 15. Procédé de production d'un polypeptide conforme à la revendication 1 ou 2, lequel procédé comporte le fait de cultiver un organisme transformé approprié, conforme à la revendication 14.
 - 16. Composition médicale comprenant un polypeptide conforme à la revendication 1 ou 2, un oligonucléotide anti-sens conforme à la revendication 6 ou 7, ou un anticorps conforme à l'une des revendications 8 à 11.
 - 17. Emploi d'un polypeptide conforme à la revendication 1 ou 2, d'un oligonucléotide anti-sens conforme à la revendication 6 ou 7, ou d'un anticorps conforme à l'une des revendications 8 à 11, en vue de la fabrication d'un médicament conçu pour le traitement d'un état associé à une augmentation pathologique de la résorption osseuse.
- 15 18. Emploi conforme à la revendication 17, pour lequel l'état associé à une augmentation pathologique de la résorption osseuse est une ostéoporose ou une hypercalcémie.
 - 19. Emploi d'un polypeptide conforme à la revendication 1 ou 2, d'un oligonucléotide anti-sens conforme à la revendication 6 ou 7, ou d'un anticorps conforme à l'une des revendications 8 à 11, en vue de la fabrication d'un médicament conçu pour prévenir la métastase de cellules cancéreuses.

įŢ,

762 707 783	862 792 804	6 20
OHGENNWAR BEBAREFEE BESEGERAGE DNOGLIVOLL BILLCLIBAG FUNDKAKTL MADEFHKKT. TWEKLACINE SRETSGBLO Heriografi band Lasysdeone stelpisavu bivilvarn iragivika aprojorisk aprojoris. Hegiografi pagnogrege biosgeppes stetappet getrassen aligiliki limivical cloteariae pppezzeage i	oahitegkgi lætaritht pedrislæg mæisfelda ravpolospo ryllpihote masgparti baspayrgag girafsejon etfadpiskt glskplf tirossi-pa krippdpset vstagbert akprrpppag pgavsss-21 gvpvyapkip motrego <u>enti etf</u> elkorgv Leeppeper pererarer d	
FEVIERNTE HATE ISAGIVINA APRO ALIGETILVI LETT	PGAVSSS-21 PASI	
galtvail Biilclibas Sipuswu Biviluabov Ptappet Ottrassoul	VPOLOSPO RVLLPLHOTE ParpagePPA	
56 Broscetros DN 14 LASVSDEQNEST 56 BLOSCEPPPS ST	O HODISHELDA RAN 27 VSTYGRERE! AND	erlisalvri pcocepany apirpaphy (prpshhayi k Kpitappipp vrctogivą gangoccpk (alkvpiok- r
A CONTRACTOR	Proprestry Indepopse D	NP TRUTTE KRI ZATOGEGGE
CHCEMENTS PSHCHROCHS PSHCCOCT IS	Legraphity Trossi-pa Pereraree	PCOOEPOUR VRPGTGTVE
HGREVGINDA MENGE HT - EVGENKA EGHÇH MENGIGINE (O MENGO	Qahhtpagi alshple Leeepepepe	Brlsalvrt Rptfapptpp
5 5 5	ā 2 5	ā <u>&</u>

240 GCG R	300 CCG R	360 CTC S	420 AGA D	480 GAA N
4 55	E 00	က် သွ	CA A	466 766
990	292	GA G	ATC	AA.
) () ()	igg A	SCA R	ATC H	TGG B
230 AGCG	290 TGCJ	350 TGG(410 AGG/	470 GCCT
AGA E	CCI	AGT V	C & A	CAC
A A	36C	CGA	AGC A	ر ا
220 230 240 ICAATGCCAGAGCGCCCGCCC MAERPAR	G	D TTA(0 200	O GAT
220 3ACA	280 GGCT(A· (340 AGCT	400 CATC	460 CCTGA L I
ည	CCT	AGG,	GAG S	AGA D
. 22	GGC A	GAG	ACA O	ర్ది జ
210 CAGG	260 270 280 290 30 CCCGCCCCCCCTGCCCTGCCCCCCCC P A R A L L L A L A G A L L A P R	320 330 340 350 36 GGGATGAGTTTGTGGGACCAGAGGGAGCTTACGAAGTGGCCAGAGCCT G M S L W D Q R G A Y E V A R A S	380 390 400 410 420 AAGGACCCTGGGATCCAGGATCATCCAGA K D P G I P G Q S I P A K D H P D) 440 450 460 470 48 GTGCAACTGCAGCTGGAGGGAAGGA. V Q L Q L E S R D L I L S L E R N
252	ີ່ວິ	GGA	ည္သ	66. В
3CA(CCT	GTG W	GAT	GCT 1
200 AGCT(260 GCGC	320 GTTT(380 CTGG	440 TGCA
2CA(20 20 A	S. S	, SC 33	ACT A
3AA(, SGC	GAT	GGA	SCA G
Q Q Q	,)))) L	3 3 6 6	CAA(o rgr v
190	250 3CCC P I	310 CCCGA(370 3AGC/ S	430 GACT
999	25 CGCGCGCC R A P	GCAGCCCC	37 CTTCTGAC L L S	43 GTGCTGA(V L T
190 200 210 240 240 240 220 230 240 CGCGGGGGCCCCCCGAGCTGCCAGGCCGCGGCGCGCGGCGGCGGCGGCGGGGGGGG	250 260 270 280 290 300 gccccccccccccccccccccccccccccccccc	310 320 340 350 360 TGCAGCCCGAGGGATGACTTTGTGGGACCAGAGGGGGCTTACGAAGTGGCCAGAGCCTC A A R G M S L W D Q R G A Y E V A R A S	370 380 390 400 410 420 CCTTCTGAGCACCCTGGGATCCCAGGATCATCCAGA L L S K D P G I P G Q S I P A K D H P D	430 440 450 460 470 480 CGTGCTGACTGCAGCTGGAGGCCGAGACCTGATCCTCAGCCTGGAAAGGAA V L T V Q L Q L E S R D L I L S L E R N

FIG.2b

540 TGT V				
	600 TGC A	660 LAAA N	720 AGC A	780 rcac T
GA'	D Q	igy E	, 22, 4	, ZZ
r T	. 95	F	SGT(N
530 ATGG1	590 TGCA/	650 TCAT(710 AACT(770 AGTC(
53 (GA1 D	59 'GTG V	65 'ATC I	ZAAA K	77 AAG K
CAA	CA7	CTJ	TAC Y	TAAC
C TCT	, GGA		S	ica1
520 FTATC Y I	580 3CATG H	640 CCGC	70([GA(760 ACAGC Q H
490 500 510 520 530 540 GAGGGACTCATTGCCAATGGCTTCACGGAGACCCATTATCTGCAAGATGGTACTGATGT E G L I A N G F T E T H Y L Q D G T D V	550 560 570 580 590 60 TCTCTCACTCGAAATCACACGGATCATTGTTACTACCATGGACATGATG S L T R N H T D H C Y Y H G H V Q G D A	610 620 630 640 650 66 GCATCAGTGCTCAGTACTTGCTCTGATCTTCAGGGGACTTATCATGTTTGAAA A S V V S L S T C S D L R G L I M F E N	670 680 690 700 710 720 AAAACGTACAGCTTAGAGCCCAAACACCCAGCGCTACAAACTCGTCCCAGG K T Y S L E P M K N T T D S Y K L V P A	STC.
JAC(ΓΤΑ(Υ	GA.	ACC T	
510 GGA(E	570 TTG1	630 CTCI S	690 AAAC N	750 GTGTG C G
CACC	ICAT H	rrg, c	3AA/	L SCT
CIT(3GA7	rac) T	M M	86G(
500 ATGG	560 ACAC	620 TCAG	680 AGCC	740 TCCA
SAA'	50 FCA(65 CT L	68 AGA(E	74 CATO
IGC A	Z Z	CAG(S	CIT	SAAC
CAT	J FCG	3GT(V	CAG(o GAC
490 ACTC	550 CACT(T 1	610 AGTG	670 STACA Y	730 CATGA M T
995 9	l L	ATC. S	T	3AG
490 500 510 520 530 540 TGAGGGACTCATTGCCAATGGTACTGATGT E G L I A N G F T E T H Y L Q D G T D V	550 560 570 580 590 600 CTCTCTCACTCGAAATCACACGGATCATTGTTACTACCATGGACATGTGCAAGGAGATGC S L T R N H T D H C Y Y H G H V Q G D A	610 620 630 640 650 660 TGCATCAGTCAGCTCTGCTCTCGAGCGGGACTTATCATGTTTGAAAA A S V V S L S T C S D L R G L I M F E N	670 680 690 710 720 TAAAACGTACAGCTTAGAGCCCAAACACCCCAGC K T Y S L E P M K N T T D S Y K L V P A	730 740 750 760 770 780 TGAGAGCATGACAACATCCAAGGGCTGTGTGGGTCACCAAGTCCAACCTCAC ESMINIQGLCGSQHNKSNLT

207	227	247	267	287
840 GAC T	860 870 880 890 900 AAGTACGTAGAGTTATTGTGGCAGACACAGAGAGTTTCAGAG K Y V E L V I V A D N R E F Q R	960 TGA D	1020 ATGA I D	1080 ACTG
GAG E	5 OV 0	S TT5% V	AA7	16 16 10
AGA R	F	CAC	¥TGG	ICT/
NAAG K	O GAG E	SAT N	IO NGTO V	70 STTS
830 CATA H K	890 AGAG/ R E	950 GCCA/ A N	1010 GAAG1 E V	1070 GAGT1 E F
AGG R	N	ATC	GTG V	CA7
800 810 820 830 840 TCCCCTGGAACCTCTCAAATGCGGGCAAGAGGCATAAGAGAGAG	GAC	GAG	980 990 1000 1010 1020 CCACTGAACATCCGGATCGTGGTGGGAATGA P L N I R I V L V G V E V W N D	1040 1050 1060 1070 108 TCTATAAGCCAGGACCCATTCACCCAGGCTCCATGAGTTTCTAGACT S I S Q D P F T R L H E F L D W
820 GCA A 1	880 GCA(940 ATA(I	1000 GGTAG	1060 CAGG(R 1
CGG	GTG V	TTA	r CTG L	ACC T
ATG M	ATT I	CGA R	SGTG V	LTTC
810 TCAA Q	870 GGTT V	930 GCAG	990 GATC I	1050 ACCCAT
TCT S	e CTO	AAG K	9 200 8	10 16 10 10
ACC	GAG E	GTT	ATC	CAG
O GGA G	O GTA V	AAÀ K	O AAC	0 AGC S
800 CCTG(P G	860 TACG1 Y V	920 GAGA/ E K	980 CTGA/ L N	1040 ATAAG I S
TCC S	AAG	CTG L	CCA	TCT
GTC V	ACC T	GAC	AGA R	.TGC
790 GAT D	850 ATG/ M :7	910 AAAG K D	970 TACA Y	1030 CAAAT K C
GAA	A.A.G.	GGA	TTT F	1 GAC D
790 800 810 820 830 840 CATGGAAGATGCCCCTCCAAAATGCGGCCAAGAAGGCATAAGAGAGACACAGAAGAAGAAGAAGAAGAAGAAGAAGAAG	850 860 870 880 890 900 CCTTAAGATGACCAAGTACGTAGGTTATTGTGGCAGACAACAGAGAGTTTCAGAG	910 920 930 940 950 960 GCAAGGAAAAGCTGAAAGCTTAATAGAGATCGCCAATCACGTTGA Q G K D L E K V K Q R L I E I A N H V D	970 980 1000 1010 1020 CAAGTTTTACAGACCACTGGATCGAGTGGTGGGAATGA K F Y R P L N I R I V L V G V E V W N D	1030 1040 1050 1060 1070 1080 CATCGACAAATGCTCTATAAGCCAGGACCCATTCACCAGGCTCCATGAGTTTCTAGACTG I D K C S I S Q D P F T R L H E F L D W

-	ロ
(1
(ک
	_
Ļ	Ļ

1190 1200 GCATGTGCACACA M C T A E Q	1250 1260 GCCGCAGTGACCTTGGC A A V T L A	1310 1320 AGAGGGCTGCAGCTG R G C S C	1330 1340 1350 1360 1370 1380 CAGAATGGCCGCAGAGAGGAGGCTGCATCATGAACCCGTCCACGGGGTTCCCATTCCC R M A A E K G G C I M N P S T G F P F P
1190 12 GCATGTGCACAGAA M C T A E	1250 12 GCGCAGTGACCTTG A A V T L	1310 13 AGAGGGCTGCAGC R G C S	1370 13 3GGTTCCCATTC 3 F P F
1190 GCATGTGCACTGCA M C T A	1250 GCGCAGTGACC A A V T	1310 AGAGGGGCTGC R G C	1370 3GGTTCCCA 3 F P
1190 SCATGTGCACT M C T	1250 GCCCAGTG A A V	1310 AGAGGGGC R G	1370 3GGTTC 3 F
.119 GCATGTGC M C	125 GCCGCA A A	131 AGAGG R	137 366
3CATG	3CC	V	
ည္ဟ		යු ස	ÄCG
S &	GGT	CTG	S. S
180 ATG	240 CTT L	SACA T	1360 AACCCGTC N P S
ATC I	CCC P	GAC	SAAC
ر اردر	S. S.	E CAJ	ZAT(
A A A	230 IGAC D	390 3AA(N	1350 'GCATC
11 'ATG	12 TTCA	12 iATC	orte.
999	CA7	9	7666 6
SATC I	D D	30 7.T.T.	03 03 03 03
116 ACC T	122 ATC	128 XAAC N	1340 3AAAG(K G
ACC	GTC V	CAC H	IGA(
	GTT V	9	A A
150 CAA	210 166A 6	1270 3CTG	1330 GGCC(A /
TTC F	SGA G	GAG E	LAT(
TAT Y	S	CA1	SAG/
	1150 1160 1170 1180 1190 1200 TTATTTCCAAGGAACCACCATCGCCATGCCCATCATGAGCATGTGCACAGGAACA Y F Q G T T I G M A P I M S M C T A E Q	TTATTTCCAAGGAACCACCATCGCCATGGCCATCATGAGCATGTGCACTGCAGAACA Y F Q G T T I G M A P I M S M C T A E Q 1210 1220 1230 1240 1250 1260 GTCTGGAGGAGTTGTCATGGACCATTCAGACAGCCCCTTGGTGCCGCAGTGACCTTGGC S G G V V M D H S D S P L G A A V T L A	TTATTTCCAAGGAACCACGCATGGCACCCATCATGAGCATGTGCACTGCAGAACA Y F Q G T T I G M A P I M S M C T A E Q 1210

FIG.2e

407	427	447	467	487
1390 1400 1410 1420 1430 1440 CATGGTGTTCAGCAGCAGGAAGGACCTGGAGGCTAGCCTGGAGAAGGGCATGGG	1500 GAAA N	1560 ATCG	1620 AGTG	. 1680
14 VTG	15 15 16A	15 AT	16 AG	36.00
75		75 X) (2)	S
	<u> </u>	CAC	පු ප	Z A
1430 GAGAA E K	90 SAA K	50 NTG	e K	S S
H GA	1490 CGGA/ R K	1550 GAATC E C	1610 GCGC/ A H	1670 TCCAC S S
 	360	3AG	8	, j
SS [999) (1)	151	99
STA S	T. G.	AC.	O TGJ	CAC R
1420 GGCT/ A S	1480 CTTT(F (1540 AGAA(E	1600 TGCT(A \	1660 ATGC/ C
E E	D V	25.5		GCA 1
اد نظ	CA(C	76C	. C V	2
10 GAC	70 AAG	30 3AC	1590 TGAAG	GA,
1410 IAGGA(1470 TCAAC	1530 GTGAC D	159 TGA	1650 CAGGA
3G.A	951	rer C	5 -) J
1400 1410 1420 1430 144 SAGCTGCAGCAGGACCTGGAGGGCCATGG S C S R K D L E A S L E K G M G	AGA E	AGA E	1580 1590 1600 1610 1620 CTACCACCTGTACTCTGAAGCCAGGATGCTGTGTGCGCCACGGGCAGTG T T C T L K P D A V C A H G Q C	1640 1650 1660 1670 1680 GTCAGCTGCAGGAACTGCAGGGGGCTCCAGCAACTCCTG
SAG S	05 CC.	E GA	0 TGT C	AAG K
1400 TGCAG C S	1460 CTACC L P	1520 GGAG/ G E	1580 ACCTG T C	1640 CTGAA L K
AGC	AC.	AG.		AGO
CAGC	TC.	AAG	CTA	3TC 0
တို့ သို့	် (၁)	0 25 H	0 99 A	ુ સું
1390 GTTCA F S	1450 CCTCT	1510 TGTGG V E	1570 TAACG N A	1630 AGAC7 D C
Y.	ည်း	TA.	်ည်	E & J
139 ATGGTGTT M V F	1450 1460 1470 1480 1490 1500 GATGTGCCTCTTCAACCTACCAGGGTCAAGCCAGGCCCTTTGGGGGCCCGGAAGTGTGGAAA M C L F N L P E V K Q A F G G R K C G N	1510 1520 1530 1540 1550 1560 TGGCTATGTGGAGAGAGAGAGGGAATGCACGAATCG G Y V E E G E E C D C G E P E E C T N R	1570 CTGCTGTAACG C C N A	1630 CTGTGAAGACT C E D C
<u>ပ</u>	3	E)	<u>ق</u> ح	5°

FIG.2f

202	527	547	267	587
1740 TACA . H	1750 1760 1770 1780 1790 1800 TGATGGCCACCCGTGGCTGGCTTACTGCTACAACGGCATCTGCCAGACCCA D G H P C Q G V D G Y C Y N G I C Q T H	1860 GCTT ; F	1870 1880 1890 1900 1910 1920 TGAGCGAGTCAACTGTGCAAAGACTCCAAGAGCGC E R V N S A G D P Y G N C G K D S K S A	1930 1940 1950 1960 1970 1980 CTTCGCCAAATGTGAGAGAGTGCCAAGTGTGGGAAATCCAGTGTCAAGGTGGTGC F A K C E L R D A K C G K I Q C Q G G A
CC TO	GAC	CTG	1 GAG S	1 1766 6
GTA Y	ČČŠ Ժ	CAT	CAA	LAGG G
1690 1700 1710 1720 1730 1740 TGACCTCCCAGGATTCTGCACGGGACTGCCCCCTCACTGTGTGTG	1750 1760 1770 1780 1790 1800 GATGGCCACCGTGTGGTTACTGCTACAACGGCATCTGCCAGACCCA D G H P C Q G V D G Y C Y N G I C Q T H	1810 1820 1830 1840 1850 1860 TGAGCAGCAGTGTCACGCTCTGGCATCTGCTT E Q Q C V T L W G P G A K P A P G I C F	1910 GACTC D S	1940 1950 1960 1970 1980 SAGCTGAGATGCCAAGTGTGGGAAATCCAGTGTCAAGGTGGTG E L R D A K C G K I Q C Q G G A
17 CAA N	17 CAT I	18 17 17 17	19 AĞA D	19 16 16 16 16
AGC A	ອ	A A	ICA K	7 22
O TCC P	CAA	ACC P	0 11GG G	30 VAA7
1720 CTGT C]	1780 CTAC/	1840 STAAAC K F	1900 \CTGTG C G	1960 3GAAAA K I
TCA H	CTG	3TG(A	STA/ N	3TG(
)) ()	Y	AGC G	VTG(NGT(
1710 CTGC(1770 ATGGT	1830 GACC/	1890 CTTA	1950 CCAAG K
GAC T	1 GGA		VTCC P	ATG(
.AGG	55.	icTC ₩	SAG/ D	GAG/
1700 TGCAC	1760 CAGGG	1820 ACGCT T L	1880 GCAGG A G	1940 CTGA(
CT CTC	17	18 ICAC	18 7TG(A	150
ATT	STS	rrg. V	S	37G
AGA E	000	0.000 CO	CAZ	30 VAT(
1690 GACCTCCCAGAAT D L P E F	1750 CCAC H	1810 GCAGT	1870 AGTC/	1930 TTCGCCAAATG F A K C
5 5.7.	TGG G	တ္လ	GCG R	کر 4
TGA	TGA D	TGA	TGA E	CT

209	627	647	299	687
2000 2010 2020 2030 2040 FCATTGGTACCAATGCTGTTTCCATAGAAACAAATATCCCACAGCAGGA I G T N A V S I E T N I P Q Q E	2050 2060 2070 2080 2090 2100 AGGAGGTCGGATTCTGCGGGGGGGGGCCCATGTGTGTGTG	2110 2120 2130 2140 2150 2160 AGGGCTTGTGCTGCAGGAAGGAAAAATCTGCCTCAATCGTCGATG G L V L A G T K C A E G K I C L N R R C	2170 2180 2190 2200 2210 2220 TCAGAATATCAGTGTTCACAAGTGTGCCATGCCAGGGCCGAGGGGT Q N I S V F G V H K C A M Q C H G R G V	2230 2240 2250 2260 2270 2280 ATGTAACAACAGGAAGAATTGCCACTGTGAAGCCCACTGGGCTCCACCCTTCTGTGACAA CON NORK NOCH CEAHWAPPPFCDK
~ 25 0	AG.	TC.	. AG	TG. D
ACA 0	GCC	7CG R	ည် 🗠	Ci
2007 P 2007	SAT M	2150 CTCAA L N	2210 CACGG H G	2270 CCCTT P F
2030 ATCCC I P	2090 IGACAT D M	21. CGT	22 CCA H	22 ACC P
N	GA1	o TGC	. 3TG	700 P
ACA		I	GCA(A GC
2020 AGAA/ E	2080 CTTG(L (2140 AAAAA K	2200 CATGCA M Q	2260 CTGGG W A
Z ATA I	FTAC Y	799		CCA.
TCC S	org V	E E	.TG	A A
2000 2010 2020 2030 2040 CATTGGTACCAATGCTGTTTCCATAGAAACAAATATCCCACAGGAGGA I G T N A V S I E T N I P Q Q E	2060 2070 2080 2090 2100 rctgtgccggggacccatgtgtacttgggtgatgacatgccagacco	30 GCA A	2190 ACAAGT I K C	2250 GTGAA
2010 .TGCTGT	20 ACC T	2130 3TGTGCA C A	CAC	27.02 CTG7
AAT	999	LAAG K	Seri V	C.A.C.
O T	0 0 8 8	2120 GGAACA G T	စ္တ ဗ္ဗ	\$0 rTG(
2000 GGTA(G T	2060 TGCCG C R	212 1664 6	2180 7TTCG(F G	2240 3AATTG
ATT	CTG	GCA A	· jg.v	SAA(K
GTC V	ATT	CTI	S S	R
2 1990 ACCT(2050 TCGG/R	2110 TGTGC	2170 NTATCA I S	2230. CAACA N R
2g 1 1 3GA	2 3GT G	CTT L	AAT N	N AAC N
FIG. 2g 1990 AAGCCGACCTGT S R P V	2050 GGAGGTCGGA1 G G R I	2110 GGGCTTGTGC7 G L V L	2170 CAGAATATCA(Q N I S	TGT
FIC	W Y	Ä	Ē,	*

3h
7
Q
FI
_

707	727	747	767	787
2340 AGGGCTT G L	2400 TGTATCT Y L	2460 AAAAGCT K L	2520 ACCACAC H T	2580 ACAGGCA R H
2330 CAGATAACC D N Q	2390 GATTTGTGG F V V	2450 CCACCATGG T M E	2510 GCCAGGCTC Q A H	2570 CCCCCAAGG P K D
2320 TCAGGCAAG R Q A	2380 TTGCTGCTG A A G	2440 ATAAAAAAA K K T	2500 CTCACCTTG	2560 ATTTCAATA(F N T
2300 2310 2320 2330 234(IGGAAGCACAGAGGTGGTCCCATCAGGCAAGCAGATAACCAGGGCT1 G S T D S G P I R Q A D N Q G L	2360 2370 2380 2390 2400 CCTGGTGATCTGCTGCTGCTGTGTTTGTGGTGTATCT L V S I L C L L A A G F V V Y L	2420 2430 2440 2450 2466 CGTTGATGCGCCTGTTCACACATAAAAAACCACCATGGAAAAGC L M R L L F T H K K T T M E K L	2480 2490 2500 2510 2520 CCTTCCCGGACACCCAGGCCTCACCACACACACACACACA	2540 2550 2560 2570 2580 GCCTGCTGATGAACCGGGCACCACTTTCAATACCCCCAAGGACAGGCA L L M N R A P H F N T P K D R H
2300 IAGCACAGAC S T D	2360 GTGAGCATC V S I	2420 ATGCGGCTG	2480 TCCCGGACA	2540 CTGATGAAC
2290 2300 2310 2320 2330 2340 GTTTGGCTTTGGAGGAAGCAGACAGAGCAGGGCTT F G F G G S T D S G P I R Q A D N Q G L	2350 2360 2370 2380 2390 2400 GACTGTAGGAATCCTGGTGAGCATCCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGTGTATCT T V G I L V S I L C L L A A G F V V Y L	2410 2420 2430 2440 2450 2460 CAAAAGGAAGACGTTGATGCTGTTCACACATAAAAAAACCACCATGGAAAAGCT K R K T L M R L L F T H K K T T M E K L	2470 2480 2500 2510 2520 AAGGTGTGTGCCCTTCCCGGACACCCAGGCCTCACCTTGGCCAGGCTCACCACACACA	2530 CCCCGGGAAAGGCCTG P G K G L
GTTTG F G	GACTG	CAAAA K R	AAGGT	9 9 9

807	827	847	867	887
2640	2700	2760	2820	2880
CCCACA	CAGTGG	AAAACC	CACTAG	CAGACC
P Q	S G	K P	T S	R P
2630	2690	2750	2810	2870
CGAGCCG	CGTGCACC	GCATTCG	TCTCGGCT	GCCCCCAT
R A V	R A P	GIR	S R L	A P I
GACGCTC D A R	2 ACCCCAC T P R	2 3CCCAGG A Q G	20 VGGACTT	26 CCCAGO
2620	2680	2740	2800	2860
GCCCTCC	CCACCAG/	SAGGCAGG	CTGAGCA	.GGGCACC
P L I	H Q 7	R Q A	L S R	G H F
2590 2600 2610 2620 2630 2640	2650 2660 2670 2680 2690 2700	2710 2720 2730 2740 2750 2760	2770 2780 2790 2800 2810 2820	2840 2850 2860 2870 2880
CTCGCTGAAATGCCAGACATGGACATCAGCAGGCCCCTCGACGCTCGAGCCGTCCACA	GCTTCAGTCACCTCAGCGCCTCTCCAGCCCCAGTGGG	CCCTGCCAGGCCCCTGCCAGTCCAGGCCCAGGGCATTCGAAAACC	CAGTCCTCCTCAGAGCCTCCTGCTGATCCACTGAGCAGGACTTCTCGGCTCACTAG	GACCCCAGGCAGGAACCTGGGCACCGCCCAGCCCCATCAGACC
S L K C Q N M D I S R P L D A R A V P Q	L Q S P Q R V L L P L H Q T P R A P S G	P A R P L P A S P A V R Q A Q G I R K P	S P P Q K P L P A D P L S R T S R L T S	T P G Q Q E P G H R P A P I R P
O ATĠGACA M D I	o srecrec V L L	SCCAGTCO	TGCCTG(GGCAGCA
2600	2660	2720	2780	2840
CAGAACA	CAGCGAG7	CTGCCCG	AAGCCTCTG	ACCCCAGG
Q N M	Q R V	L P A	K P L I	T P G
TCGCTGAAATG	2650	2710	2770	2830
	CTTCAGTCACC1	CCTGCCAGGCCC	AGTCCTCCTCAG	GCCTTGGTGAGG
	L Q S P	P A R P	S P P Q	A L V R
25. CTCGCTGA S L K	GCTTCA L Q	CCCTGC P A	CAGTCC S P	2830 TGCCTTGGTGAG A L V R
		•	•	

FIG. 2j

20	40		80	100
120 3CTAC L H	180 ragcc s P	240 3AGAG R G	300 SCTAG L D	360 rgcaa a s
TGCA(SAAGG	cress	CTGAT(L I	TACAC Y T
80 90 100 110 120 XGCGGGCGTCCCCGGTTCTGCTTGCTGCCTCTGCTAC A G V A R F C L L A L A L Q L H	130 140 150 160 170 180 TGGCCGCTGCGAGCCGGATGGACCCACAAGGTAGCTAGCC W P L A A C E P G W T T R G S Q E G S P	200 210 220 230 240 GAACTCATAATACCTCAGTGGCGGACTTCAGAAAGCCCTGGGAGAG E L I I P Q W R T S E S P G R G	260 270 280 290 300 AGAGCAGAGCTCAGGCTGATCCTAG R A E L R V M A E G R E L I L D	320 330 340 350 360 GAGCACCTTTTTGCTCCAGCCTACAGAACCTGCTACACTGCAA E H L F A P A Y T E T C Y T A S
GCTCI A. L	GGAAG G S	TCAG	IGGGCC G R	GAAA(E T
100 TGCTG	160 CAAGA	220 366ACT	280 SCTGAA	340 FACACA
CTGCT	GACCA	GTGGC	CATG(AGCCI A
90 CGGTT R F	150 GGATG	210 ACCTCA P 0	270 :AGGG1 R V	330 GCTC(A P
TCGCC	AGCCG	TAATA	AGCTC	TTTT
A 9 93999	140 GTGCG C E	200 ACTCA L I	260 AGCÁG	320 GCACC H 1
CGCGC R A	GCGGC A A	CATGA H E	CTCAG L R	AACGA
7. 3a 70 <u>ATG</u> CCCGGGC	130 CGCTG	190 CCGCTACAGC, P L Q H	250 AAGCATCCAC K H P L	310 CTGGAGAAGA L E K N
FIG. 3a 70 80 90 100 110 120 TCATGCCCGGGCGCGCGCCCCGGTTCTGCTTGCTCTCGCTCTGCTCGCTAC	130 140 150 160 170 180 ATTGGCCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGCCCAAGGTAGCCAAGAGGTAGCC W P L A A C E P G W T T R G S Q E G S P	190 200 210 220 230 240 CTCCGCTACAGCACTCATGAACTCCTCAGTGGCGGACTTCAGAAAGCCCTGGGAGAG	250 260 270 280 290 300 GAAAGCATCCACTCAGGGCCAGAGCTCAGGGTCATGGCTGAAGGGCGAGAGCTGATCCTAG	310 320 330 340 350 360 ACCTGGAGAAGGACCACCTTTTTGCTCCAGCCTACAGAAACCTGCTACACTGCAA L E K N E H L F A P A Y T E T C Y T A S
r _	44	•		

_	\supseteq
7	7
Ċ	ว
	-

120	140			200
370 380 390 400 410 420 GTGGCAATCCTCAAACCAGGCACTGAGGATCACTGCTTTTACCACGGGACTGGACTGGAATCACTGCTTTTACCACGGGACTGGACTGGAATCACTGCTTTTACCACGGGACTGGACTGGAATCACTGCTTTTACCACGGGACTGGACTGGAATCACTGAATCACTGAATCACTGAATCACTGAATCACTGAATCACTGAATCACTGAATCACTGAATCACTGAATCACTAATAATAATAATAATAATAATAATAATAATAAT		490 540 510 520 530 540 TTATAGTGAGAAGTAACCTCAGCTACCTCAGCCCAAC I V R S N L S Y I I E P V P N S D S Q H 160	550 560 570 580 590 600 ACCGTATTTACAGATCCGAACTCTCACGCTGCCCCCGGGGAACTGTGGGTTCGAGCACT R I Y R S E H L T L P P G N C G F E H S 180	640 650 660 AGTTTACACATCAGACCAAAAGCAACCTC F T H Q T K K Q P R
370 380 3 GTGGCAATCCTCAAACCAGCACGCTGAA G N P Q T S T L K	430 440 49 TGAGGGACGTGGTCCAGTGTCACC	490 500 51 TTATAGTGAGGTAACCTCAGCTACATC I V R S N L S Y I	550 560 57 ACCGTATTTACAGATCCGAACATCTCACG R I Y R S E H L T	610 620 630 CCGGCCCACCTCGAAGGACTGGGCCCTTCA G P T S K D W A L Q

FIG. 3c

220	240		280	300
720 CTG D	780 .TGG	840 FTCG	7 222 006	960 CTC
C 99 A	CA	TG.	TAC	TG(A
GGT V	SCT L	ACT L	CTC S	Z &
CT	CAA K	D TGC A	O CTA Y	O TGA D
710 TACC Y I	770 ICGC/ R	830 LATT(I	. 88 22 P	950 CATG H
GCT1 L	740 750 760 770 780 CAGAAGAATCGACATGACGGATGCCACCAAACGCTCATGG Q K N R H D Q D A T K R K L M E	800 810 820 830 840 GTTGATAAGTTTTACCGCTCCCTGAACATCCGAATTGCACTTGTCG V D K F Y R S L N I R I A L V G	860 870 880 890 900 ACGCATGGGGATAAGTGTGAAGTTTCAGAGAATCCCTACTCTACCC T H G D K C E V S E N P Y S T L	920 930 940 950 960 AGTTGGAGGCGCAAGCTGCTTGCTCAGAAGAGCCATGACAATGCTC SWRKLLAQKSHDNAQQ
E	AC T	CAT.	AGA E	SAA
700 CGT(760 TGC(820 GAA(N	880 TTC/ S	940 TCA(
STAC Y	D GAY	, F	AGT.	IGC A
SÁA(K	CCA(CTC(S	IGA/ E	r L
690 CTAT(750 ATGA(810 ACCG	37.0 3.T.G.	930 AGCT
STCJ S	ACA7	8] rta(R K	CAA(
ACA(rcg/ R	GTT	GGA'	<u>ي</u> ∞
CTZ	SAA.	J FAA(C G	O GAG
680 \GAT	740 GAAG/ K	800 FGAT/ D	86(3CA' H	920 TTGG/
B. B.	o JCA	rGT.	J	rag. S
)))	F.F.	TAT.	TG(בַּלָ
670 GAAA K	730 TATGCAGAGT1 Y A E F	790 ATTGCCAACTA I A N Y	350 36TG V	910 TGGTCCTTTCT W S F L
ATC M	GCA A	GCC A	GAGE E	S
670 680 690 700 710 720 GCAGAATGAAACGGGAAGATCTACCTCATCATGAAGTACGTGGAGCTTTACCTGGTGGCTG R M K R B D L H S M K Y V E L Y L V A D	730 740 750 760 770 780 ATTATGCAGAGTTTCAGAAGAATCGACATGACCAGGATGCCAAACGCAAGCTCATGG Y A E F Q K N R H D Q D A T K R K L M E	790 800 810 820 830 840 AGATTGCCAACTATGTTGATAAGTTTTACCGCTCCCTGAACATCCGAATTGCACTTGTCG I A N Y V D K F Y R S L N I R I A L V G	850 860 870 880 890 900 GCTTGGAGGTGTGGGGATGGGGATAAGTGTGGAGGTTTCAGAGAATCCCTACTCTACCC LEVWTHGDKCE	910 920 930 940 950 960 TCTGGTCCTTTCTTAGTTGGGGGGCGAAGCTGCTTGCTC WSFLSWRRKLLAQKSHDNAQ

÷	j
7	
C	j
H	

320	340	360	380	400
1020 ATGGCCA M A M	1080 ATTGGTG I G V	1140 rctgcac s a h	1200 3GGCACC 3 H P	1260 CAGACAG
1010 GCCCCCTCA A P L 1	1070 GAGAATGCC	1130 AGCCATGAT S H D S	1190 GCCGCCACC	1250 AGGTATCTGO R Y L O
980 990 1000 1010 1020 GCAGGTCCTTCCAAGGCACCATTGGCCTGGCCCCCTCATGGCCA	1040 1050 1060 1070 1080 ACCAGTCTGGAGGTTAGCATGGACCACTCCGAGAATGCCATTGGTG Q S G G V S M D H S E N A I G V	1100 1110 1120 1130 1140 TGGCCCATGAGTTGGCCATGATTCTGCAC	1160 1170 1180 1190 1200 AGTGCAGCCGCTGCATCATGCCCGCCCCCCCGGGCACC S A A D G G C I M A A A T G H P	1220 1230 1240 1250 1260 TGTTCAGTTGGTGTAACAGGAGGCTGGACAGGTATCTGCAGACAG FSWCNRELDRYLOTG
990	1050	1110	1170 1180	1230
AGGCACCACC	AGTTAGCATO	TGGCCACAA(TGGCGCTGCATCATC	TAACAGGAAC
G T T	V S M	G H N	G G C I M	N R K
980	1040	1100	1160	1220
GTCCTTCCA	GTCTGGAGG	CCATGAGAT	TGCAGCCGA'	CAGTTGGTG
S F Q	S G G	H E I	A A D	S W C
970	1030	1090	1150 1160 1170 1180 1190 1200	1210
AGCTAATCACGGGCAG	TGTGCTCCGTGTACCA	TAGCCTCCACTGTGGC	ACTGCTGTTGCCAGGCGATGGCGGCTGCTGTGGCCGCCGCCGCGGGCACC	CTTTCCCCAAAGTGT1
L I T G R	C S V Y Q	A S T V A	C C S A S A A D G G C I M A A A T G H P	F P K V F
AGCTAA	TGTGCI	TAGCCI	ACTGCI	CTTTCC
L 1	C S	A S		F F

٩)
~)
	٠
C)
\vdash	
L	

500	480	460	440	420
.1560 CTATT Y Y	1500 GTTC V R	1440 CATG H G	1380 ITGTA C K	1320 AGGT R C
4. FF. >	15 663 V	ECC H	ATO C	GA CA
CAA	SCA O	TGC A	95A B	့ ည ≃
O T	SGA E	GTG C	O GGA	O AGG G
1550 CCCCAC P T	1490 TCGGGA(R E	1430 AGAGT E C	1370 AGAGG E E	1310 TGGAG G G
1510 1520 1530 1540 1550 1560 GCCAATGTCACCTCCCCGAGTTCTGCACGCCAAGTCTCCCCACCTGCCCCACCTATT Q C D L P E F C T G K S P H C P T N Y Y	1450 1460 1470 1480 1490 1500 GTTCCTGCCACCAGTGCAGGTGGCTCCTGGAACCCAGTGTCGGGAGCAGGTTC S C C H Q C K L V A P G T Q C R E Q V R	1390 1400 1410 1420 1430 1440 AGAACCCTTGCTGCAACTGCACTGCACGGGGCAGAGTGTGCCCATG N P C C N A S N C T L K E G A E C A H G	1340 1350 1360 1370 1380 ACCTGGAAGACGTGAAGAGAGAGAATGTA L E D G E E C D C G E E E C K	1270 1280 1290 1300 1310 1320 GAGGAGGGATGTCTCTCCAACATGCCGGACACTAGGAGGCCGGAGGT G G M C L S N M P D T R T L Y G G R R C
CAC	CAG	9	, 55 5	ic TC
1540 CTCCC	80 ACC T	1420 AGGAAG	1360 ACTGT	1300 GGACG T
15 TCT S	1480 GGAACC G T	1:4 AAG K	13 GAC D	13 AGG R
AAG	CCT	CTG L	TGT C	ACT
360,	CTC	ACT	3.44	SAC
1530 ACCGC T G	1470 GTGGC V A	1410 TGCAC C T	1350 GAAG/ E E	1290 CCGG/ P D
SCA T	1 TGG V	ACT C	1 676 E	TGC
ည်ပ	CC	Z Z	ည်	ICA M
O GTI	CA X	o SCTC S	AGA D	CA/N
1520 CGAG1 E F	1460 CAGTGCA	1400 TGCC1 A S	1340 GGAA(E I	1280 CTCC/ S l
ည္သ	· 22	CAA	ال ال	TCT L
Z Z	E CA	C C	STA Y	3ŢĢ C
1510 CAATGTGACC	1450 TCCTGCTGCC S C C H	AACCCTTGCT	1330 GTGGCAACGGGT, G N G Y	1270 GGAT(
15 TGT C	14 767 0	13 CCT P	13 AAC N	12 666 6
CAA.	5	AAC	25.5	GGA
)55	GT	AG.	£5)	GA.

520	540		580.	009
1620 CATGTGCC M C L	1680 SCTCGATC L D L	1740 GCTTGA	1800 TGCCAGA C Q S	1860 ACCTTGA T L N
1610 CTACAACGGO Y N G	1670 SCGGCCTGCC R P A	1730 TGTGGCAA(C G K	1790 SAAGATTCAC K·I Q	1850 ACCACCATO T T I
1600 GGCCTACTG(A Y C	1660 ACCTGGAGĊC P G A	1720 CCTATGGAAAC YGN	1780 CAAGTGTGGS K C G	1840 ATCTATTGAC S I D
1590 IGGGTGGCCA G G Q	1650 GCTGTGGGG L W G	1710 TGGTGACAC G D T	1770 CAGGGATGC R D A	1830 1840 CAACGCAGTATCTAT
1580 1590 1600 1610 1620 GCACCCCTGCGAGGGTGGCCAGGCCTACAACGGCATGTGCC TPCEGGAGGGTGGCCAGGCCTACAACGGCATGTGCC	1640 1650 1660 1670 1680 CAGTGCCAGCAGCTGTGGGGACCTGGCCTGCCTCGATC Q C Q Q L W G P G A R P A L D L	1700 1710 1720 1730 1740 GEGAATGCTGCTGGTGACCTATGGAAACTGTGGCAAGGGCTTGA V N A A G D T Y G N C G K G L N	1760 1770 1780 1790 1800 AAGTGCAGTCCCAGGATGCCAGA K C S P R D A K C G K I Q C Q S	1820 1830 1840 1850 1860 CCCCTGGAATCCAACGCAGTATCTATTGACACCACCATCACCTTGA P L E S N A V S I D T T I T L N
F1G. 3t 1570 ATCAGATGGATGGC/ Q M D G 7	1630 1640 1650 1660 1670 1680 TCACTTACCAGGACGAGCCAGCCAGCCCAGGCCCTGCCCTCGATC T Y Q E Q C Q Q L W G P G A R P A L D L	1690 1700 1710 1720 1730 1740 TTTGCTTTGAGGGGGGGCTGGTGGTGGCAAGGGGCTTGA C F E R V N A A G D T Y G N C G K G L N	1750 1760 1770 1780 1790 1800 ATGGCCAATACAGGAAGTCCCAGGGATGCCAAGTGTGGSAAGATTCAGTGCCAGA G Q Y R K C S P R D A K C G K I Q C Q S	1810 1820 1830 1840 1850 1860 GCACCCAGGCCCCTGGAATCCAACGCAGTATCTATTGACACCACCATCACCTTGA T Q A R P L E S N A V S I D T T I T L N
FIC ATC	TCA(TTTC	ATGG	GCAC

١

620	640		680	200
1920 GAAG E G	1980 AACC N H	2040 GGGA G K	2100 GGCT G W	2160 TTGC L P
GGAG	CCAC	CTGT	CCCT	TCCT
1880 1890 1900 1910 1920 TCCACTGTCGGGCACCCACGTCTACCGGGGGGGGGAGGGA	1940 1950 1960 1970 1980 TGCTGGACCCAGGGCTGGTGATGACTGGAACCAAGTGTGGCCACAACC	2000 2010 2020 2030 2040 GGGCAGTGCAGGAACACCTCCTTTGAGACGGAAGGCTGTGGGA G Q C R N T S F F E T E G C G K	2090 TTGCTT C F	2150 CAGTGG S G
STCCT P	CCAAG	S IGACG	2 TCAT	2 CGAC
1900 ACCGGG	60 GGAAC G T	20 FTTGA	2080 AGAACTG N C	2140 GCAGCG1 S V
19 ICTAC	1960 rgactgg ₄ T G	2020 CTTCTT	208 CAAG	21. TGGC/
O CACG1 H V	O GTGA7 V M	0 ACCTC T: S	O AACAA N N	O GATGC D G
1890 3CACCC/ T H	1950 36CTGG1 L V	2010 GGACAC N T	2070 CAACAA	2130 3GGGAGA G D
CGGG(CCAGC P G	rGCAG C R	STCTG V C	ACCCC T P
1880 \CTGT(1940 regace D 1	2000 36CAG	2060 1CGGG(2120 3TAAC
ATCC/ I H	ATGC] M L	GAGG(E G	H 9 /2299	TTCT(F C
1870 GGCGG R	1930 GAAGGTGACA E G D M	1990 ATTTGCTTCG I C F E	2050 GCAATC	2110 2120 2130 2140 2150 2160 TCTCCACCTTTCTGTAACACCCCGGGAGATGGTGGTGGTGGTCGTTTGC S P P F C N T P G D G G S V D S G P L P
1870 ACGGGAGGCGGA G R R I	1930 GGGAAGGTGACA E G D M	1990 2000 2010 2020 2030 2040 ATATITGCTTCGAGGGGCAGGCAGGAACACCCTCCTTTTGAGACGGAAGGCTGTGGGA I C F E G Q C R N T S F F E T E G C G K	2050 2060 2070 2080 2090 2100 AAAAGTGCCACGGGGTCTGCAACAACAAGAACTGTCATTGCTTCCCTGGCT K C N G H G V C N N K N C H C F P G W	2110 2120 2130 2140 2150 2160 GGTCTCCACCTTTCTGTAACACCCCGGGAGATGGTGGTGGTCGACGAGGTGGTCGTTTGC S P P F C N T P G D G G S V D S G P L P
AC	. 	A.	A.	ŏ

FIG. 3h

		720
2220	rtggcag	, A V
2210	GGGTCCCGTGATCGCTGGGGTGTTTTCAGCTCTTCGTGTTGGCAG	F. V
2200	TTCAGCT	SAL
	GGGGTGTI	G V F S
2190	FGATCGCT	IA
2180	regereces	G P V
2170	CCCTAAGAGTGTG	> S
-	CCCCTA	P K

		740			•	160
08	TC:	_		2340	AA!	(
2280	099	CHCYRQSHKLGKPSAL		23	.TG6	RHQFSCPFRVSQSGGT
	CTO	S			TG6	ტ
0	ACC	۵,		0	GAG	S)
2270	CAA	×		2330	TCA	ð
	999	5			ATC	Ś
	ACT	ب			GGT	>
. 2260	CAA	×		2320	CAG	2
	CCA	I	•	7	CTT	ír,
	GAG	S			$\overline{\mathcal{C}}$	D .
20	ACA	ð		2310	TTG	ပ
2250	CAG	8		23	CAG	S
	CTA	>			GTT	[24
0	CTG	ပ		0	TCA	ð
2240	ICA	=		2300	GCA	H
•	3TG	ပ			305	2
	ACT.	_			GCT	٦.
2230	SCT	7		2290	CAA	×
% .	3GT	>		2	III	[L
	TTCTGGTGCTACTGTCACTGCTACAGACAGAGCCACAAACTGGGCAAACCCTCGGCTC	. _ 1			TCCCTTTCAAGCTGCGGCATCAGTTCAGTTGTCCCTTCAGGGTATCTCAGAGTGGTGGTGGAA	Д.

 . Z ď

FIG. 3i

820	840	. 098	088	006
2520 AGCT S S	2580 AGCC S R	2640 CCAC P P	00 AG	၀၀ ၁၁ -
2480 2490 2500 2510 2520 GCCACCTGCACATTGTCGCACATCTGAACAGGGCTGCTGGGAGCT P P A P L S A H L N R A G S S	2540 2550 2560 2570 2580 GCCTCGAATAGAAAGAAGGAGCCCAGGCCTCCCCCAAGCC A R I E R K E S A R R P P S R	2600 2610 2620 2630 2640 GCACCTAACTGCCCAGGACTTCTCCAGGCCTCGACCAC A P N C L L S Q D F S R P R P P	2660 2670 2680 2690 2700 CCCAGCCAATCCGGTGCCAGGCCAAAGGACCGGTCCCAGGTCAGGAG	2720 2730 2740 2750 2760 TCAGCCCCCTACTTCTGGTCCTCCAGCCCTCCAGCAGTGC Q P P T S G P Q P P R P A V P
TĠG G	ည္သ	TCG. R	GTC. S	AGC,
O TGC A	O TCC	O GCC P	O CAG	rccz
2510 GGCTC A A	2570 GCCT(P	2630 CAGG(R	2690 TCCC/	2750 GCCTCCA P P
CAG	GAG	CTC S		CAG.
GAA	CAG. R	CTT	GAC	200
2500 ATCT	2560 CAGC(2620 AGGA(2680 AAAG(2740 AGCCC
ACA, H	S S S	CCA(SCA.	27. CA(C
GGC	GGA E	GTC S	1 999	rccj
2490 TTGTC L S	2550 AGAAA R K	2610 CTACT	5 5 5 7	30 G G
24 ATT L	25 AAG R	26. CCT.	2670 3616CC V P	2730 FTCTG S G
ACC.	AGA E	C C	ر ا	rac; T
JGC.	0 AATA I	rAA(CAA	2000 P
2480 ACCT(P	2540 TCGA/	2600 ACCT	2660 AGCC/	2720 GCCC(P
် ည	الم	79. V	, <u>7</u> 00, 4	CAC
NTC(S	5	ည်	L L	
2470 TTGA	2530 CCAGAAGCTGG P E A G	2590 TGCCC	2650 AGGCA	2710 ACCTCCCTGCT T S L L
2GT.	29 NGA/ E	ZATC	26 3AA(K	27 S
2470 2480 2490 2500 2510 2520 TGCGCGTTGAATCGCCACCTGCACCATCTGAACAGGGCTGCTGGGAGCT R V E S P P A P L S A H L N R A G S S	2530 2540 2550 2560 2570 2580 CCCCAGAAGCTGGGATAGAAAGAAGGGTCAGCCAGGGGCCTCCCCCAAGCC PEAGAR IERKESARRPPPSR	2590 2600 2610 2620 2630 2640 GACCCATGCCCTACTGCCTACTGTCCCAGGACTTCTCCAGGCCTCGACCAC PMPPAPNCLLSQDFSRPPP	2650 2660 2670 2680 2690 2700 CTCAGAAGGCACTCCCAGCCCAAAGGACCGGTCCCAGGAG Q K A L P A N P V P G Q R T G P R S G G	2710 GCACCTCCCTGCT T S L L

20	GA	H	2880	> 5	ပ္ပ
2820	CAA	×	86		55
	AGCTACCCGAGTACCGATCACAGAGGGTTGGAGCAATAATTAGCTCCAAGA	LPEYRSQRVGAIISSKI			GI CGAGAAGITICITGITCCGATGGAAGACITCCGGATGCCATGGAAGGICC
0	TAG	S	c		A.16
2810	AAT	—	2870		3
	AAT	-			SAI
	AGC	A		000	3
2800	TGG	ဗ	2860		ACL
N	GGT	>	6		AAG
	GAG	8		9	155
90	ACA	O	2850		₹ 3
2790	ATC	S	28) (<u> </u>
	9	~		6	5
0	GTA	X	_) 9. e	
2780	CGA	ट्य	2840		AGI
•	ACC	Δ.			ACA
	GCT	-1			3
2770	AAA	×	2830		בופ
2	TCC	<u>م</u>	6		GAA
	CTGTTCCAA	>		5	CIAGAAGI
	ပ			E	-

58

78

98

38

18

CCCTCGCTATGGGGCCGCGCGCTCTCGCCCTTGCCTCTGCGACTAAGGTGGCTGC M I I **TCCATCTTTCTTCTTATGAAATTATTACTCCTTGGAGATTAACTAGAGAAGGGAAG** TGGCGTGTGGCTTGCTGGGCCCAGTCCTCGAGGCCGGGCGACCAGACTTGGAACAGACTG CTCTGGGGCCCAGTTCACAGATCTCTTACGTCATCCAGGCCCAAGGAAAACAGCATA TTATTCACTTGGAAAGAAACACAGACCTTTTACCTAATGATTTTGTAGTTTACACCTACG R P D L E Q T 2 ð ~ بر از 290 110 230 350 त्य 2 2 0 H ¥ R L T (T, 100 တ 160 220 340 280 ď 0 ය ¥ Z __ A ₽ ۵, 그 표 R A · L S P 90 ۵. S 150 210 270 330 **O** به دی 80 140 200 320 260 z 田 7 7 9 S 2 G တ (T) S 310 130 190 250 م S ಅ

	178	158	138	118
660 TCA	600 TCT	540 ACT	480 GAG G	420 GAG
6 TGT	6 AGT V	ACA H	CAG	TCG R
3AG	76. G	CTC.	ACT	CTA.
JCC O	590 - AGATG' R C	O CAG	0 TGG, G	JCA H
650 3CAT	59 GAG, R	530 CAAC/ N	470 CTTT(F (410 TTGTC C F
GGA	TCT	GCA(CTG	CCA ₩
3GA(3CC P	rcr L) A	GAG S
640 TGAG	580 GGA(500 510 520 530 540 TTGGAGAATGCCAGTTTTGGAATTGAACCTCTGCACAACAGCTCACACT L E N A S F G I E P L H N S S H F	440 450 460 470 480 GGAGTGCAGAATTCCGCGGTTGCTGTGAGGCGCCTGCTTTGGACTCAGAG G V Q N S A V S A C F G L R G	400 ACA(
3GA.	OCA(IGA E	rgt, V	V
999	H H	AAT' I	TGC.	CAA
630 CACA	570 GCAT	510 TTGG	450 CGGT	390 ATCC
GAC	5 TGG G	5 ITT F	CGC A	ECA H
AGG	3GA' D	CAG' S	rtc s	rga(D
3GA	CAT(. CCA	ZAA.	S
620 GAAG	560 2000 P	500 3AAT	440 3CAG/	380 ACTC
4GA(rta(Y	3GA(AGT(CCT
ZAC	NIT	rtt(L	366,	S S
610 3GGAC	550 560 570 580 590 600 GAGCACATATTTACCCCATGGATGGCCATCCAGGAGCCTCTGAGATGTGGAGTCT E H I F Y P M D G I H Q E P L R C G V S	490 GCA7 H	430 GGA(E	370 380 400 410 420 AAGGAAGGCTCCCTACTCTCTGACCATCCCAACGTACAGAGCCATTGTCACTATCGAG K E G S L L S D H P N V Q S H C H Y R G
CAG	3CA(i L	430 TATGTGGA Y V E	GGA
610 620 630 640 650 660 CTAACAGGGGACACCACAGGAGGAAGGCACACACAGGGGGGG	550 560 570 580 590 600 TTGAGCACATATTTTACCCCATGGATGGCCATCCAGGAGCCTCTGAGATGTGGAGTCT E H I F Y P M D G I H Q E P L R C G V S	490 500 510 520 530 540 GCTTGCTGCATTTGGAATTGAACCTCTGCACACAGCTCACACA L L H L E N A S F G I E P L H N S S H F	430 GCTATGTGGAG(Y V E (370 380 390 400 410 420 ACAAGGAAGGCTCCCTACTCTGACCATCCCAACGTACAGAGCCATTGTCACTATCGAG KEGSLLSDHPNVQSHCHCHYRG

Ç	ر
T	•
TI TI)
Ц	-

218	238	258	278
720	780	840	900
CTGTTCA	AGAGAAG	CGAATTG	.GGAGCTG
L · F I	R E E	R I V	G A G
710 FATGTGGAG	770 ACTGCTGTG	830 FTAAACATT	890 NTAATTGGA I I G
700	760	820	880
CAGACCCGC	CGGAACCAGA	FACATCATG7	CCTATCAATA
680 690 700 710 720	750	810	860 870 880 890 900
CGCAGAAGAGGCTGTTCTACCACAGACCCGCTATGTGGAGCTGTTCA	ATGATGGGA(GATAGCATG	STAGAAATTTGGACAGACAGAAATCCTATCAATATAGAGGGGGCTG
R R R R V L P Q T R Y V E L F I	M M G I	D S M	L E I W T D R N P I N I I G G A G
680	740	800	860
AGÄAGAGCT	AGGTACGAC	AACTACCTG	ATTTGGACA
R R A	R Y D	N Y L	I W T
670 680 690 700 710 720	730 740 750 760 770 780	790 800 810 820 830 840	850 860 870 880 890 900
CTCAGCTGCTGCAGAAGAAGAGCTGTTCTACCACAGACCCGCTATGTGGAGCTGTTCA	TTGTTGTAGACAAGGAACGACGACCAGACTGCTGAGAAGAAG	AGATGATTCGCTTAGCAACTTCGAATTG	TGCTGGTTGGACAGACAGACAGAATCCTATCAATATTGGAGGAGCTG
Q L L R R R R V L P Q T R Y V E L F I	V V D K E R Y D M M G R N Q T A V R E E	M I R L A N Y L D S M Y I M L N I R I V	L V G L E I W T D R N P I N I I G G A G
CTCAG.	TTGTT	AGATG, M	TGCTG

338	358	378	398
1080	1140	1200	1260
CACTG	TCATG	ATCCG	TAAGG
T V	H D	S G	K G
1070	1130	1190	1210 1220 1230 1240 1250 1260
FGGCCAAAT	FGGAATGAA	FTCAGGAGC	TCCAGAAACTTTAGCAGTTGCGGAGGACTTTGAGAAGTTAACGTTGAATAAGG
G Q I	G M N	S G A	S R N F S S C S A E D F E K L T L N K G
O ATGTGTT	0 ATAACCT	O FCATGAAT M N	O AGAAGTTA K L
106	1120	1180	1240
GGGATCA	TTGGGGC,	AGCTGTAT	GACTTTGA(
G I N	L G H	S C I	D F E
1050	1110	1170	1230
ACGCAGGT	CTCATGAA	JAGCAAAG	7TGCGGAG
A G	H E	A K	A E
040	100	160	1220
AGGAGCC/	ATTGTTG(TTCTGTGC	CAGTTGCAG
R S H	I V A	F C G	S C S
ratettca	TGCATCC	AGAGTGT	CTTTAGC
C S	A S	E C	F S
1030 AGGAACAG G T V	GGAGACATI E T F	115(TGATGGGA(D G R	1210 1220 1230 1240 1250 1260 GGTCCAGAACTTTAGCAGTTGCAGTGCGAGGACTTTGAGAAGTTAACGTTGAATAAGG S R N F S S C S A E D F E K L T L N K G
_	1060 1070 1080 STGGGATCAATGTGTTTGGGCAAATCACTG G I N V F G Q I T V		

7	40
r	Ġ.
1	ב
Ļ	Ļ

418	438	458	478	498
1320 CTCCTGTG S C G	1380 3TGTGAGG C E V	1440 rgcatatg a y g	1500 SAAGACCA K T S	1560 GATGTCT D V F
1310 CAGCGCGCC	1370 AGCGAAGGA(A K E	1430 TGCTGAGTG1 A E C	1490 GTGCAGAGG(C R G	1550 CTGCCCGCCA C P P
1300 ICGAAGCCTA E A Y	1360 CTGCGGCAC C G T	1420 CAAGTCATT K S F	1480 AGGCTCCAT	1540 CTCTCAGTT S Q F
1290 CGAAGCCTGA K P D	1350 GGAGTGTGA E C D	1410 TTGTAAGCT C K L	1470 CCTTCCAGG L P G	1530 1540 CAACGGTTCCTCTCA N G S S Q
1280 1290 1300 1310 1320 CTGCTTAACATCCCGAAGCCTGACAGCCCTCCTGTG L L N I P K P D E A Y S A P S C G	1340 1350 1360 1370 1380 GTGGACCCTGGAGGAGTGTGAGGG V D P G E E C D C G T A K E C E V	1400 AAGGAAGCAC	1460 ATTGCCAGTI CQF	1520 1530 1540 1550 1560 TTCCTGAGTACTGCAACGGTTCCTCTCAGTTCTGCCCCCCCC
1270 1280 1290 1300 1310 1320 GAGGAAGCTGCCTTAACATCCCGAAGCCTGACGAGCCTACAGCGCGCCCTCCTGTG G S C L L N I P K P D E A Y S A P S C G	1330 1340 1350 1360 1370 1380 GTAATAAGCTGGTGGACGGGAGGGAGGGAGGGGGGGGGG	1390 1400 1410 1420 1430 1440 TGGACCCATGCTGAGGAGCACTTGTAAGCTCAATTTGCTGAGTGTGCATATG D P C C E G S T C K L K S F A E C A Y G	1450 1460 1470 1480 1490 1500 GCGACTGTTGTAAAGATTGCCAGTTCCTTCCAGGGGAGGCCAAGACCA D C C K D C Q F L P G G S M C R G K T S	1510 1520 1530 1540 1550 1560 GTGAGTGTGCTGCAACGGTTCCTCTCAGTTCTGCCCGCCAGATGTCT E C D V P E Y C N G S S Q F C P P D V F
	•	-	•	•

7	<u> </u>	_
		•
(<u> </u>	2
	Ι	_

538	558	578	598
1680	1740	1800	1860
CCAAGAG	rcceeca	3AGAATG	3GAGGCA
R D	s	3 N V	3 G T
1670	1730	1790	1850
AAGGCTGCC	TGTGGTTTC1	CTTCAATGCC	ACACCCAGTC
K A A I	C G F S	L Q C E	T P S F
1660 TCAAAGGCT S K A	1720 TTTGGCAAC F G N	1780 TGTGGAAAG C G K]	1840 ATCATTCAG
1650	1710	1770	1830
CATCTTTGG1	AGGTGACAGA	SAACGCGCTG	AGTACCAGCT
I F G	G D R	N A L	V P A
1640 GTGTCAGGT C Q V	1700 CCATTCTAA	1760 FGCCACTGG(A T G	1820 1830 1840 1850 1860 CCGGTGTTTGGAATAGTACCAGCTATCATTCAGACACCCAGTCGAGGCA P V F G I V P A I I Q T P S R G T
1630	1690	1750	1810 1820 1830 1840 1850 1860
AATATTATGACGCGC/	TTGCTTCATTGAAG1	FGAGTACAAGAGTG	TACAGGACATGCCGGTGTTTGGAATAGTACCAGCTATCATTCAGACACCCCAGTCGAGGCA
Y Y D A Q	C F I E V	E Y K K C	Q D M P V F G I V P A I I Q T P S R G T
	CGCGCAGTCACGTCTTTGGTTCAAAGGCTAAGGCTGCCCCAAGAGAGGCTGCCCCAAGAGGCTGCCCCAAGAGGCTAAGGCTGCCCCAAGAGGCTAAGGCTGCCCCAAGAGGCTAAGGCTGCCCCAAGAGGCTAAGGCTGCCCAAGAGAGCGCAAGAGAGCGTAAGGCTGCCCAAGAGGCTAAGGCTGCCAAGAGAGCGCAAGAGGCTAAAAAAAA		

4g
G.
Ξ

618	638	658	819	869
1870 1880 1890 1900 1910 1920 CCAAATGCTGGGGTTTCCAGCTTCCAGACCCAGGGATGGTGA K C W G V D F Q L G S D V P D P G M V N 618	1930 1940 1950 1960 1970 1980 ATGAAGGCACCAAATGTGGCAAGATTTTCAGTGTGTAAATGCTT E G T K C D A G K I C R N F Q C V N A S 638	1990 2000 2010 2020 2030 2040 CTGTCCTGAATTATGACTTCAGGGAAAATGTCATGGCCATGGGGTATGTAACA V L N Y D C D I Q G K C H G H G V C N S 658	2050 2060 2070 2080 2090 2100 GCAATAAGAATTGTCACTGTGAAGGAT N K N C H C E D G W A P P H C D T K G Y 678	2110 2120 2130 2140 2150 2160 ATGGAGGAAGCGTGGGGCCGACGTATAATGCAAAGAGCCACGCACG
1870 CCAAATGCTGGGG K C W G	1930 ATGAAGGCACCAA E G T K	1990 CTGTCCTGAATTA V L N Y	2050 GCAATAAGAATTG N K N C	2110 ATGGAGGAAGCGT(G G S V

4
4
<u>ن</u>
Ĭ

718	738	758	778	798
2220 ICTTTA F I	2230 2240 2250 2260 2270 2280 AAGAGAGATGAACAAAACCTTCAGGAAGAGAGATCACAAATGTCAGATGGCA K R D E L R K T F R K K R S Q M S D G R	2290 2300 2310 2320 2330 2340 AATCAAGCAAACGTCTCTAGACCACCAGGGG N Q A N V S R Q P G D P S I S R P P G G	2400 CCCAG P G	2460 AAACA N R
o TTTCC F L	o STCAGA S D	ACCACC	9 9 99999\ (CATGG
2210 GCCATT A I F	2270 CAAATG1 Q M S	2330 TCCAGAC S R P	2390 CCACCA(P P (2450 CCTGGGC P G H
30 3CGGCT A A	30 NGATCA R S	OCTATC	CCAGA	CGCCA
2200 TTGTTGCC V A	2260 AGAAGAGA K R	2320 ATCCTAGT P S	2380 GTGTCTC(V S	2440 CCAGACCG R P
2180 2190 2200 2210 222 CTTCTTCTTCTAATCGTCCCCTTGTTGCGGCTGCCATTTTCCTCTTT F F F L I V P L V A A A I F L F	2250 TTCAGGA F R K	2310 CCAGGAG P G D	2370 GCCCCAG	2420 2430 2440 2450 2460 CCAGGGGCCCAGGTGTCTCCAGACCACCCTGGGCATGGAAACA P G G P G V S R P P G H G N R
2) AATCG1 I V	22 AACCTI T F	23 ACAGCC Q P	23 AGGGGG G	24 CCAGG P G
2180 CTTCCT F L	2240 ACGGAA. R K	2300 CTCTAGA S R	2360 ACCACCI P P	2420 AGGGGG G G
rrctro	2 GAACTA	2 IACGTC	2 CCAGA	CACCA
2170 TGGTCI	2230 GAGATC	2290 AAGCAA	2350 ATGTCT V S	2410 GTCTCCAGAC V S R P
2170 2180 2200 2210 2220 GCCTTCTGGTCTTCTTCCTAATCGTCCCCTTGTTGCGGCTGCCATTTTCCTCTTTA L L V F F F L I V P L V A A A I F L F I	2230 2240 2250 2260 2270 2280 TCAAGAGAGATGAACTACGGAAAAGCTTCAGAAGAGATCACAAATGTCAGATGGCA K R D E L R K T F R K K R S Q M S D G R	2390 2300 2310 2320 2330 2340 GAAATCAAGCAAACGTCTAGACCAGGCGGGGGGGGGGGG	2350 2370 2380 2400 GCCCAAATGTCTCCAGACCAGGGGGCCCCAG P N V S R P P G G P G V S R P P G G P G	2410 2420 2430 2440 2450 2460 GTGTCTCCAGACCACCAGGGGGCCCAGGTGTGTCTCCAGACCACCTGGGCATGGAAACA V S R P P G G P G V S R P P G H G N R

838

818

FIG. 4i

	40 TC	2640 ATGTC	TAA	0 TTT	2630	CTT.	AAT	2620 TTAG	TAT.	ÇA'A'	2610 TAGTA	26 ATA	CTG	0 CAC	2600	ST.	TAG	2590 TATA	2 TTT	2590 2600 2610 2620 2630 2640 CTCGTTATATATATTTAGAATGTC	
•	ي ا	A A	ا ا	2 2 3	ر ا	3 ≈	ا ا ا	ا ا	I	7	A S	5 5 5	4 0	S	AIC S	AAI	ZA Z	ي ع م	E C	P Q P K I S S Q G N L I P A R P A P A P	
	80	2580	Ç	0	2570	Ş	ç	2560	N F	Ę	2550	25.55	Ç	ر 0	2540	•	3	2530	2,2	. Č	-
•	۵,	۵,	Ω,	~	တ	۵.	(I,	đ	A	۵,	o	×	A	A	>	H	م	>	<u>م</u>	(T,	
	JC	ACC	229	AAG	GTC	ည္သ	GIT	GCA	TGC	225	GCA	CAA	၁၅၁	၁ဗ္ဗာ	CTA	AAC	ACC	AGT	ည္သ	GATTCCCAGTACCAACCTACGCCGCCAAGCAGCCTGCGCAGTTCCCGTCAAGGCCACCTC	
	20	2520		0	251			200	~	2490 2500 2510	90	5 7		· ص	2480			2470	.7		

20	100	150	200		250	300	350	400	450	200	550	009	029
GGAGAAGTT	ACCCGGGGGC	CCCCCAGCTC	CCGAAGCAGC	SAME	၁၁၁၅၁၅၁၅၁၅	Teececcce	TACGAAGTGG	ACAGAGCATC	AGCTGGAGAG	ATTGCCAATG	CTCTCTCACT	AAGGAGATGC	GGACTTATCA
GCCAGAGTAG CGCGCGCGC CACGCACACA CACGGGGAGG GGAGAAGTT	GCTAGACTCG CTGCTCAGCG ACCCGGGCGC	TGCGCGAGGG GGTCGCGCCA GACTCAGGGC AGTAGGACTT CCCCCAGCTC	GECECCECE TEGGATECTE CAGCECTEGE CECEGECCC CCGAAGCAGC	READING FRAME	TGCACGCCAG GCCGGCGACA ATGGCAGAGC GCCCGGCGCG GCGCGCCCC	coccecce correct secontescr sessones resonance	TGCAGCCCGA GGGATGAGTT TGTGGGACCA GAGAGGAGCT TACGAAGTGG	CCAGAGCCTC CCTTCTGAGC AAGGACCCTG GGATCCCAGG ACAGAGCATC	GTGCAACTGC	CCGAGACCTG ATCCTCAGCC TGGAAAGGAA TGAGGGACTC ATTGCCAATG	GCTTCACGGA GACCCATTAT CTGCAAGATG GTACTGATGT CTCTCTCACT	GGACATGTGC AAGGAGATGC	TGCATCAGTG GTCAGCCTCA GTACTTGCTC TGATCTCCGG GGACTTATCA
CACGCACACA	GCTAGACTCG	GACTCAGGGC	CAGCGCTGGC		ATGGCAGAGC	GGCCCTGGCT	TGTGGGACCA	AAGGACCCTG	CGTGCTGACT	TGGAAAGGAA	CTGCAAGATG	TTACTACCAT	GTACTTGCTC
ອວອວອວອວອວ	TTTTTTGAA AAAATGAAAG	GGTCGCGGCA	TGGGATGCTG		GCCGGCGACA	CCCTCCTGCT	GGGATGAGTT	CCTTCTGAGC	CCAGCCAAGG ATCATCCAGA	ATCCTCAGCC	GACCCATTAT	CGAAATCACA CGGATCATTG	GTCAGCCTCA
GCCAGAGTAG	TTTTTTGAA	TGCGCGAGGG	9 09000909 9	٠	TGCACGCCAG	ອວອວວອວວວ	TGCAGCCCGA	CCAGAGCCIC	CCAGCCAAGG	CCGAGACCTG	GCTTCACGGA	CGAAATCACA	TGCATCAGTG

1000 1050 750 800 850 900 950 1100 1250 1300 350 1150 200 400 TGTTTGAAAA TAAAACGTAC AGCTTAGAGC CAATGAAAAA CACCACTGAC AAGGGCTGTG TGGGTCACAG CATAACAAGT CCAACCTCAC CATGGAAGAT GTCTCCCCTG GAACCTCTCA AATGCGGGCA AGAAGGCATA AGAGAGAGAC CCTTAAGATG TAGAGCTGGT TATTGTGGCA GACAACAGAG AGTTTCAGAG GGAGTGGAAG TGTGGAATGA CATCGACAAA TGCTCTATAA GCCAGGACCC GAGATCGCCA CGTGCTGGTA GAGAAGATA AAGCTTCTAC TTATTTCCAA CTGCAGAACA GTCTGGAGGA GTTGTCATGG ACCATTCAGA CAGCCCCCTT GGTGCCGCAG TCGGGATGAA CCATGACACA GCAGAGAAAG GAGGCTGCAT AGCAGCTGCA TCGTCCCAGC TGAGAGCATG ACGAACATCC TCAGTGGGGT GCGATTAATA ATCACGTTGA CAAGTTTTAC AGACCACTGA ACATCCGGAT AGCATGTGCA CATGGTGTTC AAGTTAAGCA CTCCATGAGT TTCTAGACTG GGAACCACCA TCGGCATGGC ACCCATCATG TGACCTTGGC ACATGAGCTG GGCCACACT GCTGCAGCTG CAGAATGGCC CCACGACAAT GCTCAGCTTA TCCCATTCCC GACCTGGAGA TCCACGGGGT ACCAAGTACG ATTCACCAGG AGCTACAAAC GCAAGGA'AAA CATGAACCCG CTCGAAAATC CTGGAGAGGG

FIG. 5b

GCAGGÁAGGA C	CCTGGAGGCT	AGCCTGGAGA	CTGGAGGCT AGCCTGGAGA AGGGCATGGG GATGTGCCTC	GATGTGCCTC	1450
TTCAACCTAC	TTCAACCTAC CAGAGGTCAA GCAGGCCTTT	GCAGGCCTTT	GGGGCCGGA	AGTGTGGAAA	1500
TGGCTATGTG G	GAAGAGGGAG	AAGAGTGTGA	CTGCGGAGAA	CCGGAGGAAT	1550
GCACGAATCG C	CTGCTGTAAC	TGCTGTAAC GCTACCACCT	GTACTCTGAA	GCCAGATGCT	1600
GTGTGCGCGC A	ACGGGCAGTG	CGGGCAGTG CTGTGAAGAC TGTCAGCTGA	TGTCAGCTGA	AGCCTCCAGG	1650
AACTGCATGC A	AGGGCTCCA	GGGGCTCCA GCAACTCCTG TGACCTCCCA	TGACCTCCCA	GAATTCTGCA	1700
CAGGGACTGC C	CCCTCACTGT	CCAGCCAATG	TGTACCTACA	TGATGGCCAC	1750
CCGTGTCAGG	CCGTGTCAGG GCGTGGATGG	TTACTGCTAC AACGGCATCT	AACGGCATCT	GCCAGACCCA	1800
TGAGCAGCAG	TGAGCAGCAG TGTGTCACGC TCTGGGGACC AGGTGCTAAA	TCTGGGGACC	AGGTGCTAAA	cceecrccre	1850
GCATCTGCTT		TGAGCGAGTC AACTCTGCAG GAGATCCTTA	GAGATCCTTA	TGGTAACTGT	1900
GGCAAAGACT	CCAAGAGCGC	CTTCGCCAAA	TGTGAGCTGA	GAGATGCCAA	1950
GTGTGGGAAA	¥	TCCAGTGTC AAGGTGGTGC AAGCCGACCT	AAGCCGACCT	GTCATTGGTA	2000
CCAATGCTGT	(TCCATAGAA ACAAATATCC CACAGCAGGA	CACAGCAGGA	AGGAGGTCGG	2050
ATTCTGTGCC	GGGGGACCCA	TGTGTACTTG	GGTGATGACA	TGCCAGACCC	2100
AGGGCTTGTG	CTTGCAGGAA	CAAAGTGTGC	TTGCAGGAA CAAAGTGTGC AGAAGGAAAA ATCTGCCTCA	ATCTGCCTCA	2150
					•

FIG. 5d

ATCGTCGATG		AGTGTGTTC	TCAGAATATC AGTGTCTTCG GCGTTCACAA CTGTCAAA	しまずしつはむよう	000
			VUOUD I IDOD	0140001010	077
CAGTGCCACG		GCCGAGGGGT ATGTAACAAC AGGAAGAATT	AGGAAGAATT	GCCACTGTGA	225
AGCCCACTGG	GCTCCACCCT		TCTGTGACAA GTTTGGCTTT	GGAGGAAGCA	230
CAGACAGTGG	TCCCATCAGG	CAAGCAGATA	TCCCATCAGG CAAGCAGATA ACCAGGGCTT GACTGTAGGA	GACTGTAGGA	235
ATCCTGGTGA	GCATCCTGTG	GCATCCTGTG TCTGCTTGCT	GCTGGATTTG TGGTGTATCT	TGGTGTATCT	240
CAAAAGGAAG	ACGTTGATGC	GGCTGCTGTT	ACGTTGATGC GGCTGCTGTT CACACATAAA AAAACCACCA	AAAACCACCA	245
TGGAAAAGCT	AAGGTGTGTG	CACCCTTCCC	AAGGTGTGTG CACCCTTCCC GGACACCCAG TGGCCCTCAC	TGGCCCTCAC	250
CTTGGCCAGG	CTCACCACAC	CTCACCACAC CCCCGGGAAA GGCCTGCTGA	GGCCTGCTGA	TGAACCGGGC	25.5
ACCACATTTC	AATACCCCCA	AATACCCCCA AGGACAGGCA	CTCGCTGAAA TGCCAGAACA	TGCCAGAACA	260
TGGACATCAG	CAGGCCCCTC	GACGCTCGAG	CAGGCCCCTC GACGCTCGAG CCGTCCCACA GCTTCAGTCA	GCTTCAGTCA	265
CCTCAGCGAG	Tecrccrecc	TCTCCACCAG	TGCTCCTGCC TCTCCACCAG ACCCCACGTG CACCCAGTGG	CACCCAGTGG	270(
CCCTGCCAGG	CCCCTGCCCG	CCAGTCCTGC	CCCCTGCCCG CCAGTCCTGC AGTCAGGCAG GCCCAGGGCA	GCCCAGGCA	275(
TTCGAAAACC	CAGTCCTCCT	CAGAAGCCTC	CAGTCCTCCT CAGAAGCCTC TGCCTGCTGA TCCACTGAGC	TCCACTGAGC .	280(
AGGACTTCTC	GGCTCACTAG	Tecctregre	GGCTCACTAG TGCCTTGGTG AGGACCCCAG GGCAGCAGGA	GGCAGCAGGA	285(

RE	READING FRAME	ME			
ACCTGGGCAC CGCCCAGCCC CCATCAGACC TGCCCCTAAG CATCAAGTAC	CGCCCAGCCC	CCATCAGACC	TGCCCCTAAG	CATCAAGTAC	2900
CCAGACCTTC	CCACAATGCC	CCAGACCTTC CCACAATGCC TATATCAAGT GAGAAGCCAG CCCAGACCGG	GAGAAGCCAG	CCCAGACCGG	2950
TCCTCAACAG	TGAAGACAGA	TCCTCAACAG TGAAGACAGA AGTTTGCACT ATCTTCAGCT CCATTGGAGT	ATCTTCAGCT	CCATTGGAGT	3000
TGTTGTTGTA	CCAACTTTCC	TGTTGTTGTA CCAACTTTCC GAGTTTCTAA AGTGTTTAAA ACACCATTCT	AGTGTTTAAA	ACACCATTCT	3050
CTCCAGACCC	TGGAGCCACT	CTCCAGACCC TGGAGCCACT GCCATCGGTG CTGTGTGG GTGCTTTGTG	CTGTGCTGTG	GTGCTTTGTG	3100
TACTIGCTCA	GGAACTTGTA	TACTTGCTCA GGAACTTGTA AGTTATTAAT TTATGCAGAG	TTATGCAGAG	TGTCTATTAC	3150
TGCGCAGGGC	GCCGTAGCAG	TGCGCAGGGC GCCGTAGCAG GCATTTGTAC CATCACAGGG CTTTTCTACA	CATCACAGGG	CTTTCTACA	3200
GAAGGAAGGC	TCCTCGTGCT	GAAGGAAGGC TCCTCGTGCT TTTGTTTTC TGGAGGACTT GAAATACCCT	TGGAGGACTT	GAAATACCCT	3250
GCTTGATGGG	ACCTAAGATG	GCTTGATGGG ACCTAAGATG AGATGTTTAC TTTCTATTCA AGGCCTTATC	TTTCTATTCA	AGGCCTTATC	3300
GGAAAATAGC	TCCCCACCTT	GGAAAATAGC TCCCCACCTT CCCAAGGCTG TTATGGTACC AGACACAG	TTATGGTACC	AGACACACAG	3350
CTCAGGACAC	CCCAGGGAGA	CTCAGGACAC CCCAGGGAGA ACCTGGCATG GGTTTTCTTT GTTTGCTTTC	GGTTTTCTTT	GTTGCTTTC	3400
ATTTTATCTT	TTATATTTTG	ATTITATCIT TIATATITIG GIATCCCIAT CITGGGITGI AGCCAGGGCC	CTTGGGTTGT	AGCCAGGGCC	3450
TTCAGGAAGG	TCTTGGGCCA	TICAGGAAGG ICTIGGGCCA CIGCATGCIA AIGGCCTICA GGICCIGCAC	ATGGCCTTCA	GGTCCTGCAC	3500

7.0.1					
CCTGAAGCTC	TCAGACAACA	AGTAGGATCT	TCAGACAACA AGTAGGATCT GCTTTCTAGC CAGCAGCTTT	CAGCAGCTTT	355(
GGAGAGAACC	TGGGGTACTG	TGGGGTACTG AAAAGAAGGT	TTGGGGTGTG	GTTATACCAG	360
GATGGAGACT	GGAATCCTAA	TCTGGGCAAA	CATCTGACCT	TGAGCTGAGC	365(
AGCCATGAGC	ACCTCTAGGA	AGCAAGGACG	ACCTCTAGGA AGCAAGGACG GCTGAGGTGC TGCACAAGGC	TGCACAAGGC	370(
rcrecttiga	GAGCTGGCAG	GGGCTTCTCT	CTGGCTGCCC	TTTGCAGAGT	375(
GCTAGCTGGC	ATGGCATGTT	GTTTACATCG	GTTTACATCG GGAACAGTGG TGTTTCTACA	TGTTTCTACA	380
AGAAAGCCAC	TGCCTGGGCA	CTGCAGACCT	ccercrccre	CCCATTTAGA	385(
GCTAAGCAAA	TTACCACATT	GTCTTCTGGA	CTGTAATACA ATGACCCTGT	ATGACCCTGT	390(
GTTCTGACAG	ATAGAGGAGG	CTTTCTATGG	CTTTCTATGG AACCATAACT	ATTTTCANAT	395(
GTGAACTAGT	AACCAGATCT	AGTCGATCAA	CTCTGGAGAT AGAAATCTCC	AGAAATCTCC	4000
TTTTTACTGC	AAGGCTCGAC	TTATTAAAAA	TTAGGCAGAA TCCATATGCT	TCCATATGCT	405(
TGCAAAAGCT	ATAACCACGT	GGAATGCTCT	TCTCATGGCA	CAGCCTGAGT	4100
CTGGTATCCT	TATTAGTAGC	CATTGGACAA	TATTAGTAGC CATTGGACAA AGCACCCAAA GTTACCTGTG	GTTACCTGTG	4150
TGTTCTTTC	AAGGCATCCT	AATTTCTTCA	AAGGCATCCT AATTTCTTCA GCATAGAGAG ACTCGGTCTT	ACTCGGTCTT	4200
CCTCACATTC	TGAACATACC	TATCAATGAC	TGAACATACC TATCAATGAC TAAGNCAGCA AGGCAATCCG	AGGCAATCCG	425(

D					
TTTCCGAATA C1	CTGAGTTGCT	FGAGTTGCT CACGGNAAGG CAACCTCAGC CCAGGNAAAC	CAACCTCAGC	CCAGGNAAAC	4300
TTTTTCCTC	TTTTTCCTC TGNTCTTTCA	GTATGTGACT	GGGGAGCTAC CTTCAGAAGC	CTTCAGAAGC	4350
AAATTTTCAA	AAATTTTCAA GGTGGNCTCA ACCCCATNGG	ACCCCATNGG	ATGAAAGNTA	TTTTTTAAA	4400
AAATAATTAA	AAATAATTAA TGGTAATGCC AGAGGGCTTT	AGAGGGCTTT	CCTGGCNTCC	CCTGGCNTCC AGATNGGGGC	4450
GTAGGNTTGA	GTAGGNITGA CTAGCTITCA	CGACAGAAGG	TAAATGACAG CAGTCCTCTA	CAGTCCTCTA	4500
CCTCGTCTGA	CCTCGTCTGA CTGCTTTAAG ATCAAGGCTT	ATCAAGGCTT	CTTTGGAAGG GTAACTAACA	GTAACTAACA	4550
TTAATGGCTG	TTAATGGCTG GCCTGTGCCT	TGAAGCAGAA	GGGAAATAC AGATAAGGAA	AGATAAGGAA	4600
TITGGTTTGC	TTTCTAGAAT	CCAAAACTGT	ATCCAGCATT GGGAAGCATG	GGGAAGCATG	4650
GTCTTCATGA	GTCTTCATGA CTGGGTAAAT	AAATCCACGT	CACAGATGCA	TAAAAGAATA	4700
ACTCTTATGA	ACTCTTATGA CATGCCTCTT	TTTGTGGCAC AGAGACAATA	AGAGACAATA	TTGCTGCCAC	4750
TGAGATGCAT	TGAGATGCAT ACAAATTTC	TGTAACTGAT	TGTAACTGAT ATGTCATTCA	GTAGTTGTAT	4800
TAAGGCCAAA	TAAGGCCAAA CATCCACAAC	TGTAAAGACT	TATAGAGTTG	TGTGGGCGTT	4850
GTCTTGTGAG	GICTIGIGAG ACACACAAAG CCICAGCIGA AGCGIAIGAG	CCTCAGCTGA	AGCGTATGAG	crccrccrcc	4900
AGGTGGGAGT	GATGGGGAGG		CTAGAACAC ACAAAGACAA	CAGAAGAGCT	4950
TTGGTTTGGG	GGGGGTGCAG AGAGAGTGTG GTTTAGAGGA AGTTGGAGCC	AGAGAGTGTG	GTTTAGAGGA	AGTTGGAGCC	2000

.CT 5050	AT 5100	ÄG 5150	TA 5200	TG 5250	AA 5300	CT 5350	TC 5400	AC 5450	2500	GC 5550	.GA 5600	CT 5650	CA 5700	TG 5750
GATGGTGC	ATGCATTC	GGACAGAC	TTAAAGCA	AGCCAGCATG	CATCCTTC	ATCTCATG	GGAGACCA	GTGGCTGG	ATGTGACC	CCTTAAAA	CACACAGA	TTTGCGCT	AGAAACCACA	creercre
TAAGGATGCC	TIACAGAGCC ATGCATTCAT	TICCCITACA GGACAGACÄG	TAAGATACAG TAGTTGTCAA TTAAAGCATA	GGGACCCAAT	AGTITCICIC ICCCIIICIC CAICCIICAA	TTTGCAGTCC ATCTCATGCT	CTAAAGTTGG TTGAGTCATG GGAGACCATC	GCCAGGTGGC GTGGCTGGAC	TICCIGCA CIGAGGAATA GITATAGGIT AIGIGACCCC	AGTGGGAG GCGAACCTTG CAGGCATGCC CCTTAAAAGC	TACAATAG TCCTGAGTCT GTTTTCCCAG CACAGAGA	TTCAAAATAT GCATGCCGAG TTTGCGCTCT	CACATATGGG ATGACATCAC	TAAATICI ACGGGAAGAA ATCCICCIGA CIGGICICIG
CAGTGTCCAC	AGTGCTGG CTGCTTGCTT	AATCCTGAAA	TAAGATACAG	TAACAAATT	AGTTTCTCTC	ATACAGTTCA	CTAAAGTTGG	CAGTCGGT CAAGAGCCTT	CTGAGGAATA	GCGAACCTTG	TCCTGAGTCT	TTCAAAATAT	CACATATGGG	ACGGGAAGAA
ATGATCTTCT GCCATCTCCC CAGTGTCCAC TAAGGATGCC GATGGTGCCT	GCAGTGCTGG	GAACATATTT	GGAATICCIC	ACTTCAATTT	GACAGAGGGT	GTCAAAACTA	AGAGGTATGA	TCCAGTCGGT	TGTTCCTGCA	CAAGTGGGAG	CCTACAATAG	AGTTTTCCAT	TTCCAGGTTA	
ATGATCTTCT	TACCAGCTGT GC	TTCTGAATAA GA	TGTTACTAAA GG	TTTAGCAGTA AC	AGGGTTCTTT GA	ATGACAAGAC	TATACATACT AG	CCTGAGAAAG TC	GICCICCIII	ACTTCACAGG	TGGTCTCAGA	GCAACAATGC	GTGTGAGTGT	CAAGCAACAA AT

1)					
AGGAGACATT	TTTATGCCTT	CTTAACTTTA	ITTATGCCTT CTTAACTTTA TTAGGAACTC TCAGGCTGAA	TCAGGCTGAA	2800
GCTAGGGGTC	ATTGTCCCCC	AACAAATCAA	ATTGTCCCCC AACAATCAA TACAAAGCCA TCAATGNACT	TCAATGNACT	5850
CTCGAAGAAC	TGCCAAACCC	TGATCTGTGT	GAATGITCIC AGGAGCCIGI	AGGAGCCTGT	5900
GATCCCCATG	GTGCTANAAA	GAGGCTGGAG	STGCTANAAA GAGGCTGGAG CTGGGCCAAC AAGAAGGCCT	AAGAAGGCCT	5950
AAGAGTCCTC	CTGCCTCTCA	GCAGATGTTT	ACTGAGCACT CTGAGCCAGA	CTGAGCCAGA	0009
AGCACCCCGA	CAACCAGGAG	GACGATNGCT	CAACCAGGAG GACGAINGCT GGGCAGIAGG GCGCCCAGCC	GCGCCCAGCC	6050
ACTTGCAGCT	CTITCCTCTG	AGGCCCGCTT	TGTGTTTTAA TTCCCTTCTG	TTCCCTTCTG	6100
TCAGGCCCCA	ANCAGNGGAC	ANCAGNGGAC ACTGTCCTAT	AGACCTCCCT CTNAGTTTTC	CTNAGTTTTC	6150
AGACGGCCTA	AGCCATACAC	AAATGCCCCA	AGCCATACAC AAATGCCCCA GACTAAGAAA CACCAATACN	CACCAATACN	6200
TCCCAGCAGT	CCCCAAGAAC	TGGTTTTTAA	ACACTATGAC	AAGTAGAAGA	6250
GGGTGTCACA	GAGGCCATTT	TTTTTCTTTT	CTTTCCACTC ATACTGGAAC	ATACTGGAAC	6300
CTAGGTCCTC	TCTCTACACT	CCTAGTTCCT	TTACACAACT CGGCAGTGGC	CGGCAGTGGC	6350
TCCATTACAC	CAAGGACACA	GARAAACACA	CAAGGACACA GAAAACACA GGTACCGATT TGCCTTCCTC	TGCCTTCCTC	6400
TCCTGCCAAT	CACAAGTGCC	TTACTCTGAC	CAGACCCATG	ACAAAACCTC	6450
TGTCATCCAA	GAGAGCCAAC	TCTCTACCTT	GAGAGCCAAC TCTCTACCTT TGTTACTACT TCAAGCCAAT	TCAAGCCAAT	6500

6750 6700 6.650 6800 6850 0099 0069 CATATGAATG TACCATTTCC ATGCCTTTTG TGGAGTACAG ACATATAAAC CACCAGGACC ATAGCACAGA GCAACCTCCA GNACACACAC ACACACACAC CTTGAATCTA TCCCACAGCA TATCAACCCA CAGTGACCTC CCTCCCACCG CCTTGTTCTA ATTACAAGGT GAAGATGGCC ATAGAAAATC AAGTTAGCAC TGCTCAGATC CACTGGGCAG GGGGGACTCC TTGCAGGAGA CATGGGCACA GCTTTGAAAC TCTGATGAGT TAAGTCATGC TCTGGGAGCT GTGAGCCCCA GTGGTAACTG CTAACCTTCA AGGGTCACCT AAACAGTATA GTCCAACCTT TAATTACAAA ATGCTTTTGA TGCAACCTGA ATTTCCCAAT GGCACCTATT ATAAATACTT CCATT FIG. 5j

FIG. 6a GGCGGGGGC AGGCGATGTG TGATTGCGGA CAGTGAGAG GCCGTTGCTA GGCCGGGGGC AGGCGATGTG TGATTGCGGA CAGTGAGAG GCCGTTGCTA TCATGCCCGG GCGCGGGC GTCGCCCGGT TCTGCTTGCT GCTCTCGCT CTGCGTAC ATTGGCCGCT GGCGGCGTC GAGCCGAGA GAGAGCCAG GAAGCTAC ATTGGCCGCT GCGGCGGGC GTCGCCGGAT GGACCACAG AGGCGAGAG ATTGGCCGCT GCGGCGGGC GTCGCCGGAT GGACTCCCAAAACCTCC TCAGAAACCTCC ACTCGCAAAACCTCC ACTCCTAAAACCTCC ACTCAAACCTCC ACTCTTTTGCTC CAGCCTACAAAACCTGC TACACTCCAAAACCTGC TCAGGCAAAACCTGC TCAGCAAACCTCC ACTCTAAACCTC AGTGTCACCT CAGCACCTG CCGGGAATT AGAGGACTGAAAACCTC AGTGTCACACTCA TCAGCACCTG CCGGGAATT AGAGGAACCTC AGCTACATCA TCAGCCCAAAACCTC GACGCCAACCCCAACAGC GACAGCCCAACCTC ACCTAATTAATGCAAAACCTC AGTTAACACT TGAGGTACCT GGAGCTTTAC CTGGTGGCTG GCGGGCCCCAC CTCGAAGGAC TGAGCCCTTC AGTTTACACA TCAGACCAAAACCTC GCGGGCCCAC CTCGAAGGAC TGACACTCTA TGAAGTACGT GGAGCTTTAC CTGGTGGCTG ATTATGCCAGA ACGGGAAGAT CTACACTCTA TGAAGTACGT GGAGCTTTAC CTGGTGGCTG ATTATGCCAGA GTTTCAGAAG AATCGACATG ACCCAAAACCC AAGCTCATGG	09		120	180	.240	300	360	420	480	540	009	099	720	780	
AGGCAATGGC AGGGGATGTG TGATTGCGGA CAGTGAGAGG — READING FRAME GCGCGGGGC GTCGCCGGT TCTGCTTGCT GGCTCTCGCT GCGCGCGGCG GAGCCGGGAT GGACCACAG AGGAAGCCAA GCATGAACTC ATAATACCTC AGTGGCGGAC TTCAGAAGC ACTCAGAGCA GAGCTCAGGG TCATGGCTGA AGGCCGAGAG GAACGAGCAC CTTTTTGCTC CAGCCTACAC AGAAACCTGC TCAAACCAGC AGGCTGAAGT CTGAGGATCA CTGCTTTTAC GGATGAGCTC AGTGTCACGC TCAGCCCCGG GAACTGTGGG CAGATCCGAA CATCTCACGC TGCCCCGGG GAACTGTGGG CTCGAAGGAC TGGGCCCTTC AGTTTACACA TCAGACCAAA AGGGGAAGAT CTACACTCTA TGAAGTACGT GGAGCTTTAC GTTTCAGAAG AATCGACATG ACCAGAACGC GTTTCAGAAG AATCGACATG ACCAGAACGC	GCCGTTGCTA		CTGCAGCTAC	GAAGGTAGCC	CCTGGGAGAG	CTGATCCTAG	TACACTGCAA	CACGGGACTG	AGAGGACTGA	GACAGCCAAC	TTCGAGCACT	AAGCAACCTC	CTGGTGGCTG	AAGCTCATGG	
AGGCAATGGC AGGGGATGTG TGATTGCGGA — READING FRAME GCGCGCGGC GTCGCCCGGT TCTGCTTGCT GCGCGCGGC GTCGCCCGGAT GGACCAAG GCATGAACTC ATAATACCTC AGTGGCGAC ACTCAGAGCA GAGCTCAGG TCATGGCTGA GAACGAGCA CTTTTTGCTC CAGCCTACAC TCAAACCAGC AGGCTCAAGT CTGAGGATCA GGATGAGCTC AGTGTCACGC TCAGCCCCTG AAGTAACCTC AGTGTCACGC TCAGCCCCGG CTCGAAGGAC TGGCCCCTTC AGTTTACACA AGGGGAAGAT CTACACTCTA TGAAGTACGT GTTTCAGAAG AATCGACATG ACCAGGATGC	CAGTGAGAGG		GCTCTCCCT	AGGAAGCCAA	TTCAGAAAGC	AGGGCGAGAG	AGAAACCTGC	CTGCTTTTAC	CCGGGGAATT	CCCTAACAGC	GAACTGTGGG	TCAGACCAAA	GGAGCTTTAC	CACCAAACGC	
AGGCAATGGC AGGGGATGTG READING FRAME GCGCGCGGGC GTCGCCCGGT GCATGAACTC ATAATACCTC ACTCAGAGCA GAGCTCAGGG GAACGAGCA CATTTTTGCTC TCAAACCAGC AGGCTCACGC AGGTAACCTC AGTGTCACGC AGGTAACCTC AGTGTCACGC AGGTAACCTC AGTGTCACGC AGGTAACCTC AGTGTCACGC AGGTAACCTC AGCTACATCA CAGATCCGAA CATCTCACGC GTTTCAGAAG AATCGACATG	Tgattgcgga		TCTGCTTGCT	GGACCACAAG	AGTGGCGGAC	TCATGGCTGA	CAGCCTACAC	CTGAGGATCA	TCAGCACCTG	TCGAGCCCGT	TGCCCCCGGG	AGTTTACACA	TGAAGTACGT	ACCAGGATGC	
AGGCAATGGC READIN GCGCGCGGGC GCATGAACTC ACTCAGAGCA TCAAACCAGC GGATGAGCTC AAGTAACCTC AAGTAACTC AAG	AGGGGATGTG	IG FRAME	greeceegr	GAGCCGGGAT	ATAATACCTC	GAGCTCAGGG	CTTTTTGCTC	ACGCTGAAGT	AGTGTCACGC	AGCTACATCA	CATCTCACGC	reseccerre	CTACACTCTA	AATCGACATG	
	AGGCAATGGC	- READIN	299292929	cecectec	GCATGAACTC										
FIG. 6a GGCCGGGGC TCATGCCCGG ATTGCCCGCT CTCCGCTACA GAAGCATCC ACCTGGAAA GTGGCAATCC TGAGGGACGT TTATAGTGAG ACCGTATTTA CCGGGCCCAC GCGGCCCAC	FIG. 6a ccccccccc		TCATGCCCGG	ATTGGCCGCT	CTCCGCTACA	GAAAGCATCC ACTCAGAGCA	ACCTGGAGAA	GTGGCAATCC TCAAACCAGC	TCAGGGACGT GGATGAGTCC	TTATAGTGAG AAGTAACCTC	ACCGTATTTA CAGATCCGAA	CCGGCCCAC CTCGAAGGAC	GCAGAATGAA ACGGGAAGAT	ATTATGCAGA GTTTCAGAAG	

960 1080 1140 1200 1260 1320 1380 1440 1500 1560 1620 AGATTGCCAA CTATGTTGAT AAGTTTTACC GCTCCCTGAA CATCCGAATT GCACTTGTCG GCTTGGAGGT GTGGACGCAT GGGGATAAGT GTGAAGTTTC AGAGAATCCC TACTCTACCC TCTTAGTTGG AGGCGCAAGC TGCTTGCTCA GAAGAGCCAT GACAATGCTC TCACTTACCA GGAACAGTGC CAGCAGCTGT GGGGACCTGG AGCCCGGCCT GCCCTCGATC AGCTAATCAC GGGCAGGTCC TTCCAAGGCA CCACCATTGG CCTGGCCCCC CTCATGGCCA TGTGCTCCGT GTACCAGTCT GGAGGAGTTA GCATGGACCA CTCCGAGAAT GCCATTGGTG TAGCCTCCAC TGTGGCCCAT GAGATTGGCC ACAACTTTGG CATGAGCCAT GATTCTGCAC ACTGCTGTTC TGCCAGTGCA GCCGATGGCG GCTGCATCAT GGCCGCCGCC ACCGGGCACC CTTTCCCCAA AGTGTTCAGT TGGTGTAACA GGAAGGAGCT GGACAGGTAT CTGCAGACAG GTGGCAACGG GTACCTGGAA GACGGTGAAG AATGTGACTG TGGAGAAGAG GAGGAATGTA AGAACCCTTG CTGCAATGCC TCCAACTGCA CTCTGAAGGA AGGGGCAGAG TGTGCCCATG GTICCIGCIG CCACCAGIGC AAGCIGGIGG CICCIGGAAC CCAGIGICGG GAGCAGGIIC GGCAATGTGA CCTCCCCGAG TTCTGCACCG GCAAGTCTCC CCACTGCCCC ACCAACTATT ATCAGATGGA TGGCACCCCC TGCGAGGGTG GCCAGGCCTA CTGCTACAAC GGCATGTGCC GAGGAGGGAT GIGTCTCTCC AACATGCCGG ACACTAGGAC GCTGTATGGA GGCCGGAGGT TCTGGTCCTT FIG. 6b

FIG. 60

TTTGCTTTGA	GAGGGTGAAT	TTTGCTTTGA GAGGGTGAAT GCTGCTGGTG ACACCTATGG AAACTGTGGC AAGGGCTTGA	ACACCTATGG	AAACTGTGGC	AAGGGCTTGA	1740
ATGCCCAATA	CAGGAAGTGC	ATGGCCAATA CAGGAAGTGC AGTCCCAGGG ATGCCAAGTG TGGSAAGATT CAGTGCCAGA	ATGCCAAGTG	TGGSAAGATT	CAGTGCCAGA	1800
GCACCCAGGC	CCGGCCCCTG	GCACCCAGGC CCGGCCCCTG GAATCCAACG CAGTATCTAT TGACACCACC ATCACCTTGA	CAGTATCTAT	TGACACCACC	ATCACCTTGA	1860
ACGGGAGGCG	GATCCACTGT	ACGGGAGGCG GATCCACTGT CGGGGCACCC ACGTCTACCG GGGTCCTGAG GAGGAGGAAG	ACGTCTACCG	GGCTCCTGAG	GAGGAGGAAG	1920
GGGAAGGTGA	CATGCTGGAC	GGGAAGGTGA CATGCTGGAC CCAGGGCTGG TGATGACTGG AACCAAGTGT GGCCACAACC	TGATGACTGG	AACCAAGTGT	GGCCACAACC	1980
ATATTTGCTT	CGAGGGGCAG	ATATITGCIT CGAGGGCAG TGCAGGAACA CCTCCTTCIT TGAGACGGAA GGCTGTGGGA	CCTCCTTCTT	TGAGACGGAA	GGCTGTGGGA	2040
AAAAGTGCAA	TGGCCACGGG	AAAAGTGCAA TGGCCACGGG GTCTGCAACA ACAACAAGAA CTGTCATTGC TTCCCTGGCT	ACAACAAGAA	CTGTCATTGC	TTCCCTGGCT	2100
GGTCTCCACC	TTTCTGTAAC	GGTCTCCACC TTTCTGTAAC ACCCCGGGAG ATGGTGGCAG CGTCGACAGT GGTCCTTTGC	ATGGTGGCAG	CGTCGACAGT	GGTCCTTTGC	2160
CCCCTAAGAG	TGTGGGTCCC	CCCCTAAGAG TGTGGGTCC GTGATCGCTG GGGTGTTTTC AGCTCTTTC GTGTTGGCAG	GGGTGTTTTC	AGCTCTCTTC	GTGTTGGCAG	2220
TTCTGGTGCT	ACTGTGTCAC	TTCTGGTGCT ACTGTGTCAC TGCTACAGAC AGAGCCACAA ACTGGGCAAA CCCTCGGCTC	AGAGCCACAA	ACTGGGCAAA	CCCTCGGCTC	2280
TCCCTTTCAA	GCTGCGGCAT	TCCCTTTCAA GCTGCGGCAT CAGTTCAGTT GTCCCTTCAG GGTATCTCAG AGTGGTGGAA	GICCCTTCAG	GGTATCTCAG	AGTGGTGGAA	2340
CTGGCCATGC	CTGGCCATGC CAACCCAACT	TTCAAGTTGC AGACCCCCCA GGGCAAGCGA AAGGTGACTA	AGACCCCCCA	GGGCAAGCGA	AAGGTGACTA	2400
ACACCCCTGA	ATCCCTCCGG	ACACCCCTGA ATCCCTCCGG AAGCCGTCCC ACCCCCTCT CCGGCCCCCT CCAGACTACC	ACCCCCCTCT	CCGCCCCCT	CCAGACTACC	2460
TGCGCGTTGA	ATCGCCACCT	TGCGCGTTGA ATCGCCACCT GCACCATTGT CGGCACATCT GAACAGGGCT GCTGGGAGCT	CGGCACATCT	GAACAGGGCT	GCTGGGAGCT	2520
CCCCAGAAGC	TGGGGCTCGA	CCCCAGAAGC TGGGGCTCGA ATAGAAAGAA AGGAGTCAGC CAGGAGGCCT CCCCCAAGCC	AGGAGTCAGC	CAGGAGGCCT	CCCCCAAGCC	2580

FIG. 6d						
GACCCATGCC	GACCCATGCC CCCTGCACCT AACTGCCTAC TGTCCCAGGA CTTCTCCAGG CCTCGACCAC	AACTGCCTAC	TGTCCCAGGA	CTTCTCCAGG	CCTCGACCAC	2640
CTCAGAAGGC	CTCAGAAGGC ACTCCCAGCC AATCCGGTGC CAGGCCAAAG GACCGGTCCC AGGTCAGGAG	AATCCGGTGC	CAGGCCAAAG	GACCGGTCCC	AGGTCAGGAG	2700
GCACCTCCCT	GCACCTCCCT GCTTCAGCCC CCTACTTCTG GTCCTCAGCC CCCCAGGCCT CCAGCAGTGC	CCTACTTCTG	GTCCTCAGCC	CCCCAGGCCT	CCAGCAGTGC	2760
			REA	READING FRAME	ME	•
CTGTTCCAAA	CTGTTCCAAA GCTACCCGAG TACCGATCAC AGAGGGTTGG AGCAATAATT AGCTCCAAGA	TACCGATCAC	AGAGGGTTGG	AGCAATAATT	AGCTCCAAGA	2820
		:	•		•	
TCTAGAAGTG	TCTAGAAGTG TCGAGAAGTT TCTTGTTCCG ATGGAAGACT CCGGATGCCA TGGAAGGTCC	TCTTGTTCCG	ATGGAAGACT	CCGGATGCCA	TGGAAGGTCC	2880
AGAAGAAAGA	AGAAGAAAGA CGCCTTCTCA CCCATCCTGA AGCTTTGGCA GCCTTCTGGA ACGTCCCTCA	CCCATCCTGA	AGCTTTGGCA	GCCTTCTGGA	ACGTCCCTCA	2940
TCCCCAGAAT	TCCCCAGAAT CTCCCTTCTT		ACCCGAGTGC CTCCTGCTTC CTCCGAGGCC CAGGGGGACT	CTCCGAGGCC	CAGGGGGACT	3000
CATATCCAAT	CATATCCAAT GGCTCCTAAG TGTTTGTCCT GTGCAATATA CAGCCCAGGG AGGGAAGGGA	TGTTTGTCCT	GTGCAATATA	CAGCCCAGGG	AGGGAAGGGA	3060
AGCACGGCGA	AGCACGGCGA GGAGGGTGGG AAAGGTTCTC CCTCAGCCCA CTAGCCAAGA GCTACCAGCG	AAAGGTTCTC	CCTCAGCCCA	CTAGCCAAGA	GCTACCAGCG	3120
. •						

FIG. 6e						
ATGCTCAGGG	AAGGCTTGAG	ATGCTCAGGG AAGGCTTGAG CTGGGGTCCT CCTCTGCGGA GCTTGGAGAA GGTACCCATC	CCTCTGCGGA	GCTTGGAGAA	GGTACCCATC	3180
CTGGTCCTAT	GCTGGCAGGA	CTGGTCCTAT GCTGGCAGGA ACACACGCGA GTGTCACTGA TTGGCCTCCT TCTGGGATCC	GTGTCACTGA	TIGGCCTCCT	TCTGGGATCC	3240
CAGGCTGCTG	AGGAAGCTAC	CAGGCTGCTG AGGAAGCTAC TGCTACATCC CTACCCCAAG GGGCTTGGTC AAGGTGCCTG	CTACCCCAAG	GGGCTTGGTC	AAGGTGCCTG	3300
TYCCTGGCTC	TCTGGCTGCA	TYCCTGGCTC TCTGGCTGCA TGTAATAAGC CATGCTCCCC TCCCCTGCCT TTCTTCACAT	CATGCTCCCC	TCCCCTGCCT	TTCTTCACAT	3360
TCCCACTCCC	ATATTTACAC	TCCCACTCCC ATATTTACAC GGGTCACTCT GACTCAGACA GGTACTATTT GTAAGTAGCA	GACTCAGACA	GGTACTATTT	GTAAGTAGCA	3420
TAGACAGCAG	GGGGGTGGGG	TAGACAGCÁG GGGGGGGGG TGGTCAACCT GTGTCCCCTC TGAGCGTTA TGCCAAAGGT	GIGTCCCCTC	TGAGCCGTTA	TGCCAAAGGT	3480
CACTAAGGAG	ATTTAGAATC	CACTAAGGAG ATTTAGAATC CCCATCCATCCATCCATCCATCCATCCATTCA	CATCCATCCA	TCCATCCATC	CATCCATTCA	3540
TCCATCCCCA	GTGTTCCATG	TCCATCCCCA GIGITCCATG TGTCACCITC ICCITITCCA GCATCCCIAT CCTAIGGIGC	TCCTTTTCCA	GCATCCCTAT	CCTATGGTGC	3600
TTTGGTGGTG	AACTATGGCA	TTTGGTGGTG AACTATGGCA GTCCTGACTT GCTGATGACC ATATGCTGGT	GCTGATGACC	ATATGCTGGT	GACCTACAAA	3660
TCGGGATCCT	GCCATATGGG	TCGGGATCCT GCCATATGGG GTCGCCACTG GACTTTCTGC ACTGGTTCTC AAGAGCGTTG	GACTTTCTGC	ACTGGTTCTC	AAGAGCGTTG	3720
AGCCGAGTGG	GCGTGTATGT	AGCCGAGTGG GCGTGTATGT TTGTGTGTGT GTGTGTGT GTGTGTGT GTGTGTGTGT	GTGTGTGTGT	GTGTGTGTGT	GTGTGTGT	3780
GTGTGTGTGT	GTGTGTGT GTGTGTGT	GTGAAAGAGA	CAGAGGCAAT	GTGAAAGAGA CAGAGGCAAT GAGAGACA GACATGCAGG	GACATGCAGG	3840
CAGGCCGACA	CAGGCCGACA GCTCTGCATG	TACTTGTGTT	TTACGGCCTC	TTACGCCCTC AAGCAGTATA AGGGACCTCC	AGGGACCTCC	3900
TCCTTATTTC	TCCTTATTIC TGACTCATAT	CTAAGTAAGG	TTCCCCAGGA	TTCCCCAGGA CMAGCCACAG CTGTACTGAG	CTGTACTGAG	3960
GGGGCTGAC	ATGTTTGGCA	GGGGGCTGAC ATGTTTGGCA TCCTGGCTAT AGTATTGTAT ACACAGGGCC ACCAGCCCCG	AGTATTGTAT	ACACAGGGCC	ACCAGCCCCG	4020

FIG. 6f

4080	4200	4260	4320	4380	4440	4500	4560	4620	4680	4740	4800	4860	4920
TCAACAGCCA TTATCACGCA	TGCTGAGAAG	GAACAGTTCT	AGGTGCTAAT	TGTTAGCTCT	GAGGCCTCCT	GACAGACGTG	CCGTGCTCTT	TGCTTTACCA	CCTCCACTTT	ATCACTTCTG	TGTGANGTGT	CACACAACTT	CCCTTCCCAT
CCCTAGTGGT CAGCTCTGAG GGGGGACTGG TGACTCTGAA CAGATCGATG TCAACAGCCA TGGTGAACCA GATCTGGGCA GGGTTCCCCA AACTCTATTC AACCAGAGTT TTATCACGCA	NCTCATCGGG TCTCTCCTGG TTGCTGCCCC GAGGTGATCG TCATGGAAAA TGCTGAGAAG	GG ACCTICICIT GCITGGTGCT CCGCTATITG GAACAGIICI	TACACATITG CIGGGCCIGG CCICIGAGAG GCCAICTICC ACCCCCAGAA AGGIGCIAAI	GGCACTGCAG AGGGCTCTCT AGGGGCCTCC CCGCCCCAAC AGCAAGCAGT TGTTAGCTCT	TGGAACCCTC CAGAGGAAGA GGCAAGCGTT TGACTTCCCC TTTACCACCT GAGGCCTCCT	TATATCTCTT CCCAGAGTAA GCTTTGGGAT TGTAGACATG TGGGAGCTAT GACAGACGTG	GCCTGGGGTA GAAAGATCTC AGGAAAGCAC CTTTCTCCTT TTCAGGGTGA CCGTGCTCTT	CACACTCTCT GAGGCCTCAG TCCATGTCCT ATATCAGTTT CTCTTTTGTG TGCTTTACCA	AGTGGCCGGT GACTACAGGC CACCCCGATT CTCACCACAA AGTTAGAAAC CCTCCACTTT	CTGTCCCTTG AACCATATCA GAAAAAGACC CATTTCCTTG CTCTTTGGTA ATCACTTCTG	TITITICITC TICATTACTG TGCTACCACC TCCATCCCAT GACATTATTC TGTGANGTGT	AAGAGGACGG TGTTTNTTA NTCTTGGGAG ANATGTCGGC AGCTGCTCTA CACACATT	CACTCAAGGC TTTGTCTCCA GAGGCCAGCT AGGCTGTCAC AGGCAGGAAT CCCTTCCCAT
TGACTCTGAA AACTCTATTC	GAGGTGATCG	GCTTGGTGCT	GCCATCTTCC	CCGCCCCAAC	TGACTTCCCC	TGTAGACATG	CTTTCTCCTT	ATATCAGTTT	CTCACCACAA	CATTTCCTTG	TCCATCCCAT	ANATGTCGGC	AGGCTGTCAC
GGGGACTGG GGGTTCCCCA	TIGCIGCCC	ACCTICTCT	CCTCTGAGAG	AGGGCCTCC	GGCAAGCGTT	GCTTTGGGAT	AGGAAAGCAC	TCCATGTCCT	CACCCCGATT	GAAAAAGACC	TGCTACCACC	NTCTTGGGAG	GAGGCCAGCT
CAGCTCTGAG GATCTGGGCA	rcrcrcree	GATGGGGTGG	creeeccree	AGGGCTCTCT	CAGAGGAAGA	CCCAGAGTAA	GAAAGATCTC	GAGGCCTCAG	GACTACÁGGC	AACCATATCA	TTCATTACTG	TGTTTTTTA	TTTGTCTCCA
CCCTAGTGGT TGGTGAACCA	NCTCATCGGG	GTGGGAATGG GATGGGGT	TACACATTTG	GGCACTGCAG	TGGAACCCTC	TATATCTCTT	GCCTGGGGTA	CACACTCTCT	AGTGGCCGGT	CTGTCCCTTG	TTTTTTCTTC	AAGAGGACGG	CACTCAAGGC

1.1O.08		•				
CTGCTTTGTG	AAGGGTCCCA	CTGCTTTGTG AAGGGTCCCA TACAGGTGTA TCTAGACTTC AAGGACAGGG TTTGTCTCAC	TCTAGACTTC	AAGGACAGGG	TTTGTCTCAC	4980
AGGATTGTCA	CTTAGGAGAT	AGGATTGTCA CTTAGGAGAT GAAAGAATAT TACCACATGA GGAGGAGGGG CAGTTGCAAC	TACCACATGA	GGAGGAGGGG	CAGTTGCAAC	5040
AGAACACTTT	GGTCTTCCTA	AGAACACTTT GGTCTTCCTA CACCAAGTCT GTGAGGGCAT CCAAGACTGA ATGAAAGCGC	GTGAGGGCAT	CCAAGACTGA	ATGAAAGCGC	5100
TTTTCTTATG	CATACAATGT	TTTTCTTATG CATACAATGT GAGCAAGAAC AAGAACTGTT TAAGGCACCT CTGTTCCCAG	AAGAACTGTT	TAAGGCACCT	CTGTTCCCAG	5160
CCACTGAAGA	GAGACGTCAG	CCACTGAAGA GAGACGTCAG AAGATGTTAG AATAGGTCAA AACCAAGGCT CTGGTGGACT	AATAGGTCAA	AACCAAGGCT	CTGGTGGACT	5220
GAGGGAAGGT	TTGTAGCTGC	GAGGGAAGGT TIGTAGCTGC GTTTAGTGGT ATACATCTTT AGTCCCAGCA TAGGCAGGTG	ATACATCTTT	AGTCCCAGCA	TAGGCAGGTG	5280
AATCTCGAGT	TTGAAGCTAG	AATCTCGAGT TTGAAGCTAG CCTGGTCTAA AAAGGAAGTT CCAAGACTGC CAGGGCCACA	AAAGGAAGTT	CCAAGACTGC	CAGGGCCACA	5340
CAGAGGAAAA	AAAAAACCC	CAGAGGAAAA AAAAAACCC TCTAGAAAAA CAAAAATGAA GACAGGTTCT CATGTATCGT	CAAAAATGAA	GACAGGTTCT	CATGTATCGT	5400
AGATTGGCCT TTAAGTCA	TTAAGTCACT	CT TTACCAAGGA TGATCTTTGA ACTCCTGAGT ACAGACTGCG	TGATCTTTGA	ACTCCTGAGT	ACAGACTGCG	5460
GGTGTGTGCT	ACCATGCTTT	GGTGTGTGCT ACCATGCTTT ATGTGGCCCT GGGTTCAAAC ACAGCCCTTC ATATGTATAT	GGGTTCAAAC	ACAGCCCTTC	ATATGTATAT	5520
AGCCAAACAC	TCTACAACTG	AGCCAAACAC TCTACAACTG AGCTACATCC TCCAGCCTAG GCTGTAAATG TTTTTTGGAG	TCCAGCCTAG	GCTGTAAATG	TTTTTGGAG	5580
CTAGATTAGC	TGCCTGCCAA	CTAGATTAGC TGCCTGCCAA CCTTAGAACT GCAAAGCCAT TCCTGACCTG TAAACCTCAG	GCAAAGCCAT	TCCTGACCTG	TAAACCTCAG	5640
CTCTCCATCT	CTATAAGAGG	CTCTCCATCT CTATAAGAGG TATAGCCTGG GCTAATACCG TCCAAGTTAC AACTCCTTGC	GCTAATACCG	TCCAAGTTAC	AACTCCTTGC	5700
TTGCTTTCTG	TTCCTTCTAG	TTGCTTTCTG TTCCTTCTAG CCTTGGTGAC TTCCACCAGG AAGAGAATAC CCCCTCTCTA	TTCCACCAGG	AAGAGAATAC	CCCCTCTCTA	2160
CCCCTGCTCC AAGACACT	AAGACACTGT	GT AGATGCTAGT GTCGGAGTGT TCTCTGTAAC GCGACAGTTC	GTCGGAGTGT	TCTCTGTAAC	GCGACAGTTC	5820

5940 0909 6180 6240 6352 6120 6300 0009 TITITIGITG IIGITGITAT IITITICICC ATTTGGGAGA TTGTTCACAG AGTTATTTTG TCAAATGCAT GTAATGAACA GACCCGAAGG AATCCTCCAC ACACAAGCCA GGGAACACCA ACTGGAAAGG TACCCCGTCC CAGGGAAGCC TGCTAGGGAG AGGTTCTGTA GAATCCGAGC CTAGCACCCC AAAGTCATGC ACCCAGTATC CTCTTGTATG ACTGTATATG TACATCTATG CTTCTGTTGC AATAGCCCCC CTGCAACACT GCAATAATCC TTCAGTGTCT CCCCTGGGCT CAATTCACTT CCTTATTTGA CAAAGTGGAG GTGAGACTTG TATTCTTAAA ATTGGAGGCT ITGTTTGTAT TTAATTAAAA CAAATTGTCA TGAGGAAAAA AAAAAAAAA AA ACAATATTTG ICTATGTCTG GGATCCAGGG CAAATGTGAA TTTCCTTTTG GAAGTAGICC ICCCCICICA IGICCICCIA IIGAIIGIII CAAAATACIT GAATGGGCCA TGGTGCCTTG

FIG. 6h

							•							
٠	CCTGCCGAAG		AGGTGGCTGC	GAACAGACTG	AGAAGGGAAG	AAACAGCATA	TACACCTACG	CACTATCGAG	GGACTCAGAG	AGCTCACACT	TGTGGAGTCT	CCGAGTGTCA	GAGCTGTTCA	GTGAGAGAAG
	GTTGCAAGGA TGACCGAAGN NCGGAGGCGG CGGCCGCGCG TTGAGCGGAA CCTGCCGAAG		OCCIOCCIAT GGGGCGGGG GGGCTCTOGC CCCITGCCTC TCTGCGACTA AGGTGGCTGC	TEGOSTGTEG CITECTEGEC CCAGICCTCG AGGCCGGCCG ACCAGACTTG GAACAGACTG	TCCATCITIC ITCTIATGAA ATTATIACIC CITGGAGAIT AACTAGAGAA AGAAGGGAAG	CICTGGGGC CAGITCACAG CAGAICICII ACGICAICCA GGCCCAAGGA AAACAGCAIA	TIATICACTI GGAAAGAAAC ACAGACCTIT TACCIAATGA TITIGIAGII TACACCTACG	ACAAGGAAGG CTCCCTACTC TCTGACCATC CCAACGTACA GAGCCATTGT CACTATCGAG	GCTATGTGGA GGGAGTGCAG AATTCCGCGG TTGCTGTGAG CGCCTGCTTT GGACTCAGAG	GCTTGCTGCA TTTGGAGAAT GCCAGTTTTG GAATTGAACC TCTGCACAAC AGCTCACACT	TIGAGCACAT ATTITACCCC ATGGATGGCA TCCACCAGGA GCCTCTGAGA TGTGGAGTCT	CTAACAGGGA CACAGAGGAAG GAAGGCACAC AGGGGGATGA GGAGGAGCAT CCGAGTGTCA	CTCAGCTGCT GCGCAGAAGA AGAGCTGTTC TACCACAGAC CCGCTATGTG GAGCTGTTCA	TTGTTGTAGA CAAGGAAAGG TACGACATGA TGGGACGGAA CCAGACTGCT GTGAGAGAAG
	5000000000	AME	cccrreccrc	AGGCCGGGCG	CTTGGAGATT	ACGTCATCCA	TACCTAATGA	CCAACGTACA	TTGCTGTGAG	GAATTGAACC	TCCACCAGGA	AGGGGGATGA	TACCACAGAC	TGGGACGGAA
	NCGCAGGCGG	READING FRAME	GOGCTCTCGC	CCAGTCCTCG	ATTATTACTC	CAGATCTCTT	ACAGACCTTT	TCTGACCATC	AATTCCGCGG	GCCAGTTTTG	ATGGATGGCA	GAAGGCACAC	AGAGCTGTTC	TACGACATGA
	TGACCGAAGN	REA	2929229999	CTTGCTGGGC	TTCTTATGAA	CAGTTCACAG	GGAAAGAAAC	CTCCCTACTC	GGGAGTGCAG	TTTGGAGAAT	ATTTTACCCC	CACAGAGAAG	GCGCAGAAGA	CAAGGAAAGG
FIG. 7a	GTTGCAAGGA	L	CCCTCGCTAT	TEGCETETEG	TCCATCTTTC	CTCTGGGGCC	TTATTCACTT	ACAAGGAAGG	GCTATGTGGA	GCTTGCTGCA	TTGAGCACAT	CTAACAGGGA	CTCAGCTGCT	TTGTTGTAGA

FIG. 7b

840	006	096	1020	1080	1140	1200	. 1260	1320	1380	1440	1500	1560	1620	1680
ATTCGAATTG	GGAGGAGCTG	CGTCGGAGAC	ATGGCGTTTG	CANATCACTG	ATGAATCATG	GGAGCATCCG	TTGAATAAGG	cccrccrcrc	GAGTGTGAGG	TGTGCATATG	GGGAAGACCA	CCAGATGTCT	GGCATGTGCC	GCCCCAAGAG
AGATGATTCG CTTAGCAAAC TACCTGGATA GCATGTACAT CATGTTAAAC ATTCGAATTG	TGCTGGTTGG ACTAGAAATT TGGACAGACA GAAATCCTAT CAATATAATT GGAGGAGCTG	GTTCAGTGGC GGGAAAGTT CCTTATAACT CGTCGGAGAC	ACGACAGTGC ACAGTTGGTT TTGAAGAAAG GCTTTGGTGG AACTGCAGGA ATGGCGTTTG	TAGGAACAGT ATGTTCAAGG AGCCACGCAG GTGGGATCAA TGTGTTTGGG CAAATCACTG	TGGAGACATT TGCATCCATT GTTGCTCATG AATTGGGGCA TAACCTTGGA ATGAATCATG	ATGATGGGAG AGAGTGTTTC TGTGGAGCAA AGAGCTGTAT CATGAATTCA GGAGCATCCG	GGTCCAGAAA CTTTAGCAGT TGCAGTGCGG AGGACTTTGA GAAGTTAACG TTGAATAAGG	GAGGAAGCTG CCTGCTTAAC ATCCCGAAGC CTGACGAAGC CTACAGCGCG CCCTCCTGTG	GTAATAAGCT GGTGGACCCT GGAGAGGAGT GTGACTGCGG CACAGCGAAG GAGTGTGAGG	TGGACCCATG CTGTGAAGGA AGCACTTGTA AGCTCAAGTC ATTTGCTGAG TGTGCATATG	GCGACTGTTG TAAAGATTGC CAGTTCCTTC CAGGAGGCTC CATGTGCAGA GGGAAGACCA	GTGAGTGTGA TGTTCCTGAG TACTGCAACG GTTCCTCTCA GTTCTGCCCG CCAGATGTCT	TCATTCAGAA TGGATATCCT TGCCAGAACA GCAAAGCCTA CTGCTACAAT GGCATGTGCC	GGCTAAGGCT
GCATGTACAT	GAAATCCTAT	GGGAAAAGTT	GCTTTGGTGG	GTGGGATCAA	AATTGGGGCA	AGAGCTGTAT	AGGACTTTGA	CTGACGAAGC	GTGACTGCGG	AGCTCAAGTC	CAGGAGGCTC	GTTCCTCTCA	GCAAAGCCTA	TTGGTTCAAA
TACCTGGATA	TGGACAGACA		TTGAAGAAAG	AGCCACGCAG	GTTGCTCATG	TGTGGAGCAA	TGCAGTGCGG	ATCCCGAAGC	GGAGAGGAGT	AGCACTTGTA	CAGTTCCTTC	TACTGCAACG	TGCCAGAACA	CAGGTCATCT
CTTAGCAAAC	ACTAGAAATT	GAGATGTGCT GGGCAACTTT	ACAGTTGGTT	ATGTTCAAGG	TGCATCCATT	AGAGTGTTTC	CTTTAGCAGT	CCTGCTTAAC	GGTGGACCCT	CTGTGAAGGA	TAAAGATTGC	TGTTCCTGAG	TGGATATCCT	CGCGCAGTGT
AGATGATTCG	TGCTGGTTGG	GAGATGTGCT	ACGACAGTGC	TAGGAACAGT	TGGAGACATT	ATGATGGGAG	GGTCCAGAAA	GAGGAAGCTG	GTAATAAGCT	TGGACCCATG	GCGACTGTTG	GTGAGTGTGA	TCATTCAGAA	AATATTATGA CGCGCAGTGT CAGGTCATCT TTGGTTCAAA GGCTAAGGCT GCCCCAAGAG

2220 2040 1980 2280 1800 1860 1920 2100 2160 2340 2400 2460 2520 AAACGTCTCT AGACAGCCAG GAGATCCTAG TATCTCCAGA CCACCAGGG TCTCAGGGAA ACTTGATTCC GGCTCGGCCC GCTCCTGCAC TGAAGTCAAT TCTAAAGGTG ACAGATTTGG CAACTGTGGT TTCTCCGGCA GAAGIGIGCC ACTGGGAACG CGCTGTGTGG AAAGCTTCAA IGCGAGAATG GGGTGTGGAT TTCCAGCTTG GTTCCGACGT TCCAGACCCA GGGATGGTGA TTATGACTGT GACATTCAGG GAAAATGTCA TGGCCATGGG GTATGTAACA CGTGGACAGC GGGCCGACGT ATAATGCAAA GAGCACAGCA CTGAGGGACG CITCITCITC CTAAICGICC CCCIIGITGC GGCIGCCAIT IICCICITIA TGAACTACGG AAAACCTTCA GGAAGAAGAG ATCACAAATG TCAGATGGCA CTCCAGACCA CCAGGGGCC CAGGTGTCTC CAGACCACCA GGGGGCCCAG ACCACCAGGG GGCCCAGGTG TCTCCAGACC GCCACCTGGG CATGGAAACA ACCAACCTAC GCGCCAAGC AGCCTGCGCA GTTCCCGTCA AGGCCACCTC TCAGACACCC AGTCGAGGCA TIGICACIGI GAAGAIGGCI GGGCICCCCC ACACTGIGAC ACCAAAGGAI CAAATGTGAT GCTGGCAAGA TTTGCAGGAA TTTTCAGTGT GTAAATGCTT GCCGGTGTTT GGAATAGTAC CAGCTATCAT CACCACAACC GAAAATATCT ATTGCTTCAT CTGTCCTGAA GCAATAAGAA ATGGAGGAAG **TCAAGAGAGA** GAAATCAAGC GTGTCTCCAG CCAAATGCTG ATGAAGGCAC GCCCAAATGT GATTCCCAGT GTGAGTACAA GCCTTCTGGT TACAGGACAT FIG. 7c

FIG. 7d READING FRAME ——

CICCITIAIA TAGCICCCIC ACCIGATAGI AGAATATIAG AATCITATIT ITTAAAIGIC TICAGGGAAC IGAGCAAAIG ITIGIIGITI ITITITIICCI GAIGITITICI IGAAAAGCCI
TTCTCTTCCA ACCATGAATG AACACAAACC ACCACAAAC AAGCTTTATT AACACAGGAG
CCTAGTGGGG ATTGCGAAAC ACAGGAATGT GCAGGCGCTC CGGGGGGTGT AAAGTGAACG
TITCCATOGI TAGAATGITI TCTCTGGCCA TITGTGGATI TAATGCACIT GACGTGGATI
AAGTTATICI GAGCAIGTIA CIGIAAIGAI ICICAAATIA ACIGIAITAG IGIAAGCIIT
GICACTAIGC GCIAAACGIA AICCIGACIT ITIGACCCCA GITACCAITA AIAGITICIG
GTTGACCATT TGAACATGTA TTAACTTAGG AAGACTAATT GCCAATAAGG TCTGCATTTT
CATCTIGCAT GGATTAACAG CCATTTATAT GGACTTATGT CTCTTAATGC ACAAAGAAGC
AGATATCTCG AAGGAGCTTA CACAAGAACC ACAATTACTA GATCATGATA TACTTGGAAA
GIGIGAAATA IGGIGIGIAC ICAGITAIIG GCIICCAIII IIWAIGAICI IICAACIAIA
ACAATTATGA TAGAAATCGA TTTAACACAA TCAGTTATGG GCTTCCATTT TCAAATATCT
TITCAACTGT AATGACTATG ACAGGAACTG ATTCAACTCT CAATITITGT TATGCATCAT

3720

3780

3840

3900

3931

TTGAAATTCT AAAAAAAA AAAAAAAA

3660

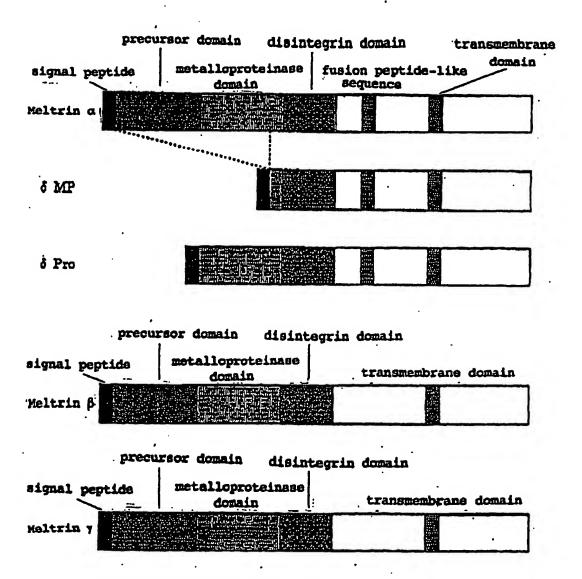
3600

3540

3480

TTGTGTATAT ATACATATA AAATAAAAC ATTTACAACA AATAAAATAC IGAGGTGTTA CCAAACCACT IGAGAATTCA IGAGCACTTT AACTCTAAAC ICTGAATTTC GIGAAGICCT CIAGAAIGII TACAITIACI AAGGIGIGCI GGGICCIGIC TCTTTTGACT AATATTTTCG TAAACATTAG GCTGGAGAAA GGAAGGAAGC AGTGGTTTCC TTAGATAACT ACAGAATTAT ACTGGTCTCT GGGATTACTC TCTCAGCTGT ATTAAAATGA TGAAAGGAAT GATATTGACA CTAAAATTTT AAACATTTAA ATTTTTTCAT AAAGAAGTIT AATAATAGGT ATATTAACTG AATTICATIA GITTITTAAA GGTAAAGCAT IGCAGCAGIG ITGITITGII IGAAGIGCAC ACICIAIGGI ACGAGGIGII PAGTATACCC AAGCAGATAG GTGTCGATCG AACAGGAGCA GGGAGAATAC TTCCAACAGT ATTTGTACTT AATCTTTCAT AAAGCTTGAT ATAATATTGT FIG. 7e

FIG. 8



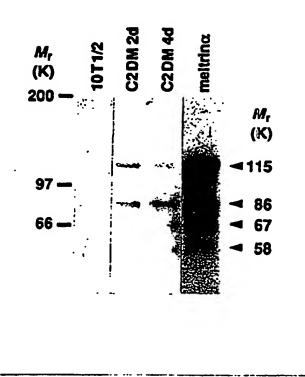


FIG. 9

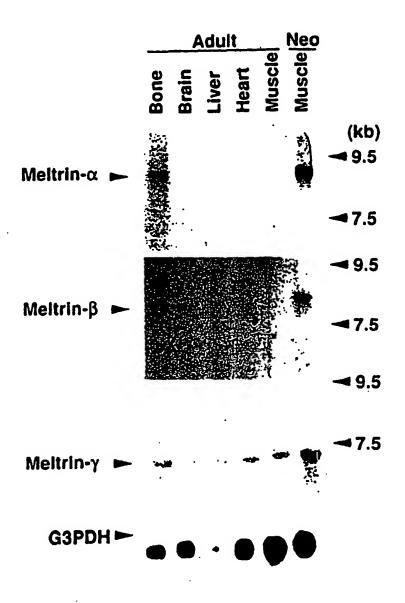


FIG. 10

FIG. 11a

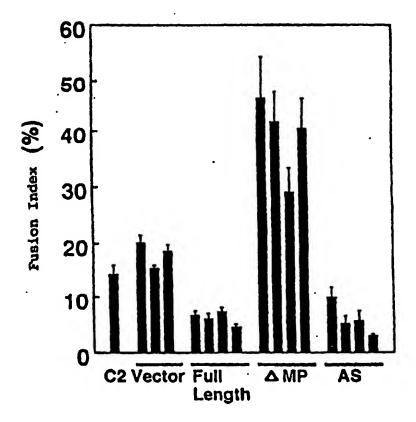
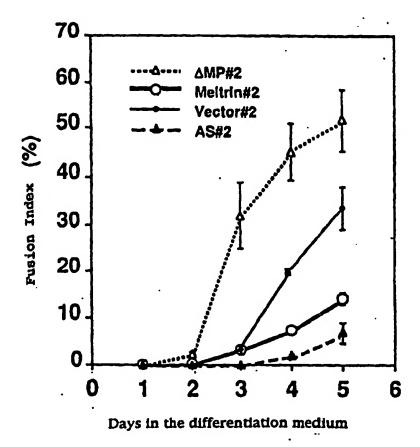


FIG. 11b



	3	40	•	9		80		1.00
AAGCCTGCAGGAACAGCGTGCAGGACTCCAGCAACTCCTGTGACCTCCAGAGTTCTGC	ACAGGGGGCACCTGCCAGCCAACGTGTACCTGCATGGGCATCTACTCAG	0	130 140 150 160 170 180 180 170 180	S L	190 200 210 220 230 240 CTCTGGGACCAGGTGCTAAACCTGGCATCTGGATTTGAGAGAGA	A	300	GEPYGN CGKVSKAGICIUGAAGIICIIIGULAAAIGUGAGAIG
TT(ر ا ا	ပ) I J	5 >	Ĺ.	S	٠٠	E E
GAG	ع · ال	S	TO T	မှ သ	TAAT	z	Ç	3 0
20 CCA	110 GCAC	H	170		230 AGT		290	X X
FIG. 12a 10 20 30 40 50 6 AAGCCTGCAGGAACAGGGACTCCAGGCACTCTGTGACCTCCCAGAGTTCTG	ר הכיני	H C P A N V Y L H D G H S C Q	1	C X N G I C Q T H E Q Q C V T	2 AGA	A K P A P G I C F E R V N S A	2 5	۲ کا در
3AC	7 Y		. 7	Б. Ш	SA S	(r)	<u>ئ</u> و	- E
CGT(78.7) ITT((7.	رز	3 _
40 CCT	100 100		160		220 GCT		280	
ACT	, J		774]	TCI		Ş	35
CA CCA	ַבָּרָי בָּי		יניני		ÇÇA		چ د	5 ~
30 CCA	90	>	150		210 CCTG	5	270 CTCT	IVI S
ACT	ָר בְּי	Z	וֹ עַלְּי	S H	8 C	<u>Д</u>	2 2	ያ ሕ የ
7995	ر الا	¥	Y.T.C.	် ပ	7.	*	۲	Ž ⊼
CAC CAC	ر ا	<u>_</u>	NA A	Z	ACC	Ы	ر ر	2 2 2
20 GTG	ر 98 م	ပ	140	×	200 TAA	×	260	ئ ئارى
AGC	r T	=		် ပ	TGC	¥	2	§ 2
AAC	ر ک	3 24	A T	5 >-	AGG	ဗ	ر ا	5 5 5
10 AGG	70 780 780	T G A S P	130	D V D G Y	190 GACC	T W G P G	250	Y X
Z 16C, 16C,	,) ,	Z V	1,	2	15.05	ဗ	2 2	کے ح
	ָלָהָ טִּ), T		7.	E	Ş	2 E3
FIG. 12a AAGCCTGCA K P A	A CA		CAT	0	CTC	7	رزر	5 5

FIG. 12b 310 320 AGAGATGCTAAATGCGGCAAG R D A K C G K 10

100	80		09	40	20	
TGTAAGCTTAAATCATTTGCTGAGTGTGCATATGGTGACTGTTGTAAAGACTGTCGGTTC	GAGTGTGTGTGCTACTCCAAAGGAATGTGAATTGGACCCTTGCTGCGAAGGAAG	240	130 140 150 160 170 180 AAGCCTGATGAGGTGCTCCCTGTGGTAATAAGTTGGTGGACGCTGGGGAA K P D E A Y S A P S C G N K L V D A G E	GCAGAGGACTTTGAGAACTTTAATAAAGGAGGAAACTGCCTTCTTAATATTCCA A E D F E K L T L N K G G N C L L N I P	GCAAAGAGCTGCATCATTCAGGAGCATCGGGTTCCAGAAACTTTAGCAGTTGCAGT A K S C I M N S G A S G S R N F S S C S	60 4GT
260 270 280 290 30 TCATTTGCTGAGTGTGGTGACTGTTGTAAAGACTGTCGGT	5		99	E	2	25
GT(AS C		25	ATA 	GT1	GTI
ACT	AAG G	0	O ACG A	TTA	3CA S	S. A.
290 AGA(වූ ස	230	17(GG/ D	177 1-	TA(S	50 TAG
TAA	ဋ္ဌာ		GGT	ည်	CTT F	CTI
rTG	GGTACTCCAAAGGAATGTAATTGGACCCTTGCTGCGAAGGAAG		140 150 160 170 18 GCCTATAGTGCTCCCTGTGGTAATAAGTTGGTGGACGCTGGGG A Y S A P S C G N K L V D A G E	GAGAAGTTAACTTTAAATAAAGGAGGAAACTGCCTTCTTAATATT E K L T L N K G G N C L L N I	ATCATGAATTCAGGAGCATCGGGTTCCAGAAACTTTAGCAGTTGCA(I M N S G A S G S R N F S S C S	AA
0 E	ည် ႕	0	O AAC K	AAO	AGA R	40 CAGA
280 GACT(GAC D	220	16 AAT N	GGAAA	S	4 TC
GT	TTG.		GT.	GA(i i	GT
ATC	AA1		GT .	AAG	99	99
270 GCAT	FIG B	210	150 TCCT S C	ATA K	S	30 :CAT
2; TG(AT(8	SCT S	AA. N	A A	38 84 84 84 84 84 84 84 84 84 84 84 84 84
GTG	ධි ස) P	TTT	A ශිරි	AGG
lGA	¥A K		GC.	IAC T	S	TC
260 TGC	ر بر کر	200	140 TAG1 S	TIT	AA'N	20 AA
TTT	ACT T	8	TAT Y	AAG K	A TG	ATG
ŢĢ.	3GT		2005	3AG	77.	TC.
AA.	STS (AAC	TT	, GC/	_ <i>(</i> 25
250 TTA	ACI	190	130 ATG/	ACTI F	GCI	10 GCTG
160	ere D		TG O	05) O	NGA S	SA S
250 TGTAAGCTTAA	gagtgtgactg E C D C		AAGCCTGATGA	GCAGAGGACT A E D F	GCAAAGAGCTG A K S C	AAA
5 7	R R		AA K	GC A	SC A	ည

120	140	160	180	200
FIG. 13b 310 320 330 340 350 360 CTTCCAGGAGGTACTTTATGCCGAGGAAAACCAGTGGTGTGTGT	370 380 390 400 410 420 AATGGTTCTCTCAGTTCTGCCAGATGTTTTTATTCAGAATGGATATCCTTGCCAG N G S S Q F C Q P D V F I Q N G Y P C Q	440 450 460 470 480 ITTGCTACAACGGCATGTGCCAGTTCTATGATGTCAAGTC C Y N G M C Q Y Y D A Q C Q V	490 500 510 520 530 540 ATCTTTGGCTCAAAGGCTGCCCCCAAAGATTGTTTCATTGAAGTGAATTCTAAA I F G S K A K A A P K D C F I E V N S K	550 560 570 580 590 600 GGTGACAGATTTGGGTTTCTCTGGCAATGAATACAAGAAGTGTGCCACTGGG G D R F G N C G F S G N E Y K K C A T G
TAC Y	. 160 0	CAA	TCT	ACT T
GAG.	ָ֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖֖	rer	AAT	200
350 TCCA(P 1	380 390 400 410 42 STTCTGTCAGCCAGATGTTTTTATTCAGAATGGATATCCTTGCCA F C Q P D V F I Q N G Y P C Q	470 TCAA	500 510 520 530 54 GCCAAGGCTGCCCCCAAAGATTGTTTCATTGAAGTGAATTCTAA A K A A P K D C F I E V N S K	560 570 580 590 60 AATTGTGGTTTCTCTGGCAATGAATACAAGAAGTGTGCCACTGG N C G F S G N E Y K K C A T G
35. TTC	41 GA1	47 CTC	53 AA(55 AG1
ATG V	ATG	ATG A	TTG	AGA K
TG. D	\GA N	VTG.	ica'	ICA.
340 AGT(400 TTCA	460 'ATT/	520 GTT	580 AATA
3 TGA	4 TAT I	GTA	TIG	TGA E
CAG S	TTI A	Y O	AGA D	CAA N
) I	rGT V	3.TG	CAA.	, j
330 AAAA K 1	390 GAT D	450 ATG	510 CCC P	570 TCTG
	CCA	9	GCC	TTC
CGA R	CAG 2	AAC	3CT	GGT
320 ATGC	380 CTGT(440 CTAC	500 CAAG K /	560 TTGT(
32 TAT	35 T.C.	4. 25.	200	56 AT7
CTT	AGT	ATT	AAG	25 25
FIG. 13b 310 CTCCAGGAGGTAC L P G G T	370 AATGGTTCTTCTCA N G S S Q	430 AATAACAAAGCCTA N K A Y	490 ATCTTTGGCTCAAA I F G S K	550 GGTGACAGATTTGG G D R F G
310 310 6466	370 rctt(430 IAAGC	490 GCT(550 GAT
FIG. 13b 310 ctrccaggag L P G G	S	CAA 4	75 a	CAG R
		L AA	CIT	rga D
FI CTI	N N	AA.	AT(I	199

300	280	260	240	220
850 860 870 880 890 900 CAGAAAAAGTGTCATGGGGTATGTAATAGCAATAGAATTGTCACTGTGAAAAT Q K K C H G H G V C N S N K N C H C E N	AAGATCTGTAGAAACTTCCAGTGTGGTGCTTCTGTTCTG	730 740 750 760 770 780 CTAGGATCAGATCCTGGGATGGTTAACGAAGGCACAAAATGTGGTGCTGGA L G S D V P D P G M V N E G T K C G A G	670 680 690 700 710 720 GTGCCTGCTATTATTCAAACGCCTAGTCGAGGCACCCAAATGTTGGGGTGTGGATTTCCAG V P A I I Q T P S R G T K C W G V D F Q	610 620 630 640 650 660 AATGCTTTGTGGGAAGCTTCAGTGTGGAATGTACCTGTATTTGGAATT N A L C G K L Q C E N V Q E I P V F G I
IGA E	GA.	IGC.	FTT	1992 1992
STG.	STG C	ည်	iga) D	ITTI F
.890 GTCA(830 TGAC D	770 ATG1 C	680 690 700 710 72 TCAAACGCCTAGTCGAGGCACCCAAATGTTGGGGTGTGGATTTCCA	650 TGTA V
TTG	rta' Y	AKA	999	P P
GAA'	GAA.	CAC	rtg(GAT/ I
880 ATAA I K	820 FTCT	760 3AAGG	700 AATG	640 AAGA(
CAA 8	AGT.	CGA.	CAA.	ACA.
TAG	TTC	TAA	CAC	TGT. V
IAA	o TGC	O GGT V) AGG G	SAA.
870 ATGT	810 AGAT D	750 GATG	690 ICGA R	630 FGAG/ E
GGT, V	IGT. V	1 66	ragʻ	STG
TGG G	GTG C	7CC P	ري ا	CA O
860 ACA'	800 CCA(740 AGA' D	680 AAC	620 GCT L
TGG G	CTT	JCC P	TCA Q	AAA
TCA H	AAA	TGT V	ľAľ I	1 66.
850 AGTG	790 GTAG	730 CAGA	GTGCCTGCTATTAN	610 TGTG
8 AAA K	CTG C	ATC.	6 TGC	6 ITT L
GAA	GAT	AGG.	SCC.	TGC.
Šø	A A	CT	· GT(N N

320

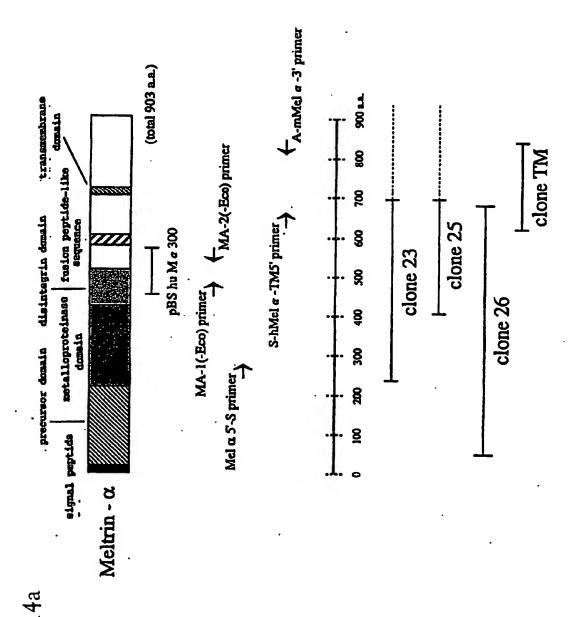


FIG. 14b

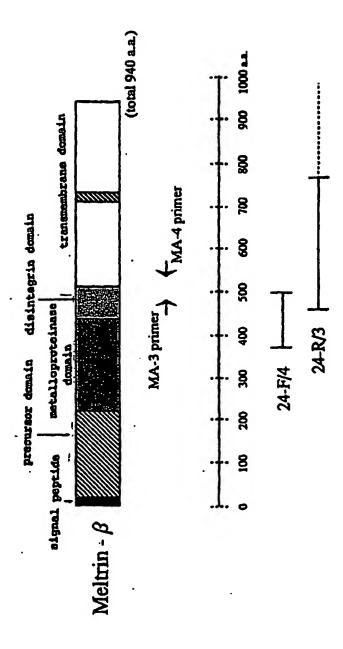


FIG. 15a

60	120	180	240	300	360 120	420
GGGGACCTCTGGATCCCAGTGAAGACTTCGACTCCAAGAATCATCCAGAAGTGCTGAAT G D L W I P V K S F D S K N H P E V L N	ATTCGACTACAACGGAAAGCAAGATCTGAAATTCTGAAATGAAGGTCTC I R L Q R E S K E L I I N L E R N E G L	ATTGCCAGCAGTTTCACGGAAACCCACTATCTGCAAGACGGTACTGATGTCTCCCTCGCT	CGAAATTACACGGTCACTGTTACTACCATGGACATGTACGGGGATATTCTGATTGAGCA R n y t g h c y y h g h v r g y s d s a	GTCAGTCTCAGCACGTGTCTCAGGGGACTTATTGGGTTTGAAATGAAAGCTAT VSLSTCSGLRGLIGFENESY	GTCTTAGAACCAATGAAAAGTGCAACCAACAAACTCTTCCCAGCGAAGAAGCTG V L B P M K S A T N R Y K L F P A K K L	AAAAGCGTCCGGGGATCATGTGGATCACATCACAACCAAACAAGAAT K S V R G S C G S H H N T P N L A A K N
); 1	55	T L	rtc. S	NAG(S	3AA K	K
CTC V	GA/	S	GA1 D	GA/	K K	ζος/ •
GAA	N	GTC V	TCT S	N N	V A	A A
P P	AGA R	GAT D	TAT Y	GAA E	P CCA	CTC L
CAT	E GAA	ACT T	S G G	TTI	TTC F	AAC N
AAT N	CTG	GGT	990; 8	999	CTC L	P P
'AAG K	AAT N	GAC	GTA V	'ATT I	AAA K	ACA
TCC S	ATA I	CAA 0	CAT	CTT L	TAC Y	AAC
GAC	ATC I	CTG L	GGA G	GGA	AGA R	CAC
TTC F	CTG L	TAT Y	CAT H	AGG R	AAC N	CAT H
AGC.	GAA	CAC	TAC Y	CTC L	ACC T	TCA S
AAG.	AAA	J.	FAC.	199 9	GCA K	GGA.
3TG/	AGC,	E GAA	ret	rcT S	AGT	TGT C
CAC	AA.	901	YC.	reT.	AAA.	rcA'
TC	99	TC); }	1931	TG/	GA7
66A 1	AAC R	GTT F	990	GCA T	CAA	055
TCT	TAC Q	S S	ACA T	TCA	AAC	TCC . R
ACC	GAC L	S	ATT Y	GTC L	TAG E	۸ 909
099 O	TTC R	TTG A	GAA N	TCA	TCT L	AAA S
ق ك	₩ ₩	A	Ŭ ≃	(b) >	5 >	××

FIG. 15k

GTGTTTCCACCCCCTCAGACATGGCCAAGAAGGCATAAAAGAGAGCCTCAAGGCA V F P P S Q T W A R H K R E T L K A ACTAAGTATGTGGAGCTGGTGATCGTGGCAACCGAGATTTCAGAGGCAAGGAAA T K Y V E L V I V A D N R E F Q R Q G K GATCTGGAAAAGTTAAATAGAGATTGCTAATCAGGTTATAC D L B K V K Q R L I E I A N H V D K F Y AGACCACTGAACATTCGGATGGTGTGGTGGCGTGGAATGGCAAA R P L N I R I V L V G V E V W N D W D K TGCTCTGTAAGTCAGCACCCATTCACCAGCCTCCATGAATTTCTGGACTGGAGATG C S V S Q D P F T S L H B F L D W R K M AAGCTTCTACCTCCCAAATCCCATGACATGCCAGCTTGTCAGTGGGGAGATG K L L P R K S H D N A Q L V S G V Y F Q GGGACCACTGGCATGGCCCAATCATGAGCATTGCAGGCACCAGTTGGGGGA GGGACCACTGGCCAATCATGAGCATTGCAGGCACCAGTTGGGGGA GGGACCACTGGCCAATCATGAGCATTGCAGGCACCAGTTGGGGGA GGGACCACTGGCCAATCATGAGCATTGCAGGCACCAGTTGGGGGA GGGACCACTGGCCAATCATGAGCATTGCAGGCACCAGTTGGGGGA GT T I G M A P I W S M C T A D Q S G G	160	540	600	660	720	780	840 280
GTGTTTCCACCACC V F P P P ACTAAGTATGTGGA T K Y V E GATCTGGAAAAGT D L E K V AGACCACTGAACAT R P L N I TGCTCTGTAAGTCA C S V S Q K L L P R GGGACCACCGGG GGT T I G							
GTGTTTCCACO V F P P ACTAAGTATG T K Y V GATCTGGAAA D L E K AGACCACTGA R P L N TGCTCTGTAAC C S V S K L L P GGGACCACCA GGGACCACCA G T I	CACC	rgga(E	AAGT7 V	I I	TCA(TCGC R	ງ
GTGTTT V F ACTAAG T K GATCTG AGACCA R P TGCTCT C S TGCTCT K L K L GGGACC GGGACC	CCAC(TATG1 Y	GAAA! B K	CTGAA L N	STAAG VS	CTACC L P	ACCAT
O C C C C C C C C C C C C C C C C C C C	IGTTT F	CTAAG K	TCTG	ACCA(CTCT(GCTTC L	GACC/ T
	ح و <u>َ</u>	AC T	6	A R	2 C	A K	<u>ဗ</u>

300	960 320	1020 340	1080 360	1140	1200	1260
ATTGTCATGGACCATTCAGACAATCCCCTTGGTGCAGCCGTGACCCTGGCCACATGAGCTG	GGCCACAATTTCGGGATGAATCATGACACTGGACAGGGGCTGTAGCTGTCAAATGGCG G H N F G M N H D T L D R G C S C Q M A	GTTGAGAAGGAGGCTGCATCATGAACGCTTCCACCGGTACCCATTTCCCATGGTGTTC V E K G G C I M N A S T G Y P F P M V F	AGCAGTTGCAGCAGGACTTGGAGACCAGCCTGGAGAAAGGAATGGGGGTGTGCCTG S S C S R K D L E T S L E K G M G V C L	TTTAACCTGCGGAAGTCAGGGAGTCTTTCGGGGGCCAGAAGTGTGGGAACAGATTTGTG F N L P E V R E S F G G Q K C G N R F V	GAAGAAGGAGGGGTGTGGGGGGGGCCAGGGAATGTATGAATCGCTGCTGCAAT E E G E E C D C G E P E E C M N R C C N	GCCACCACCTGTACCCTGAAGCCGGACGCTGTGCGCCTGTGTGTG
IGA B	NAT(SGT(V)TG C	VIII F)))	GA/
¥Ç¥.	CA Q	CAT()GT(R R	C C	.TGT.
) V	CTG) -	၁၅၅	A A G	چ در	TGC
CCT	rage S	F	VAT(ງ	AA1 N	icto '
3AC T	CTG) 	99)	3TG C	ľA T(999
CGT()))	STA(Y	SAA/	SAAC	VIGI C	NCA1
AGC A	CAG R)))	GGA(E	CCAC	B. B.	/35C •
TGC A	QGA	CAC	CCT	9999	AGA(316
766 6	ACT	TTC	CAG S	<u>မိ</u>) - -	TGT(
CCT	CAC	CGC A	GAC	III	E GGA	ည်
700 P	TGA D	GAA	GGA	GTC	766 6	GGA(
CAA	TCA H	CAŢ	CTT	SGA B	CTG.	ည္
AGA D	GAA	CAT	QÇA D	CAG.	TGA	GAA
TTC	GAT	CTG C	GAA	AGT.	GTG	<u>.</u> کر
ATTGTCATGGACCATTCAGACAATCCCCTTGGTGCAGCGTGACCCTGGCACATGAGC)	GGCCACAATTTCGGGATGAATCATGACACACTGGACAGGGGCTGTAGCTGTCAAATG G H N F G M N H D T L D R G C S C Q M	AGG G	CAG R	GGA E	GGA. E	GCCACCACCTGTACCCTGAAGCCGGACGCTGTGTGTGCGCTGTGTGTG
GGA	TTT F	AGG G	cAG S	ე ၂	AGA. E	CTG
CAT M	CAA	GAA	TTG	້າ	AGG, G	CAC
TGT V	CCA #	TGA	CAG S	TAA	AGA	CAC
AT I	<u> </u>	GT K	AG S	T.	S B	Ü V

TGCCAGCTGAAGCCTGCAGGAACAGCGTGCAGGACTCCAGCAACTCCTGTGACCTCCCACCCA	<u>ئ</u>
ACCTC L L L G ATGGG G Q AGCAG R A A A A A A A A A A A A A A A A A A	ت ت
ACC ATG	ပ
TG/CG/CG/CG/CG/CG/CG/CG/CG/CG/CG/CG/CG/CG	ပ
CTG R H L C C C C C C C C C C C C C C C C C C	0
CTCC S CCTC T T CTGC C C C C C C C C C C C C C	0
CAA Y Y CCA O O GAA GAA GG	_
S S CGT CCT CCT CCT CCT CCT CCT CCT CCT CCT	۵.
CCCC PP CCCC PP V V CATCCATCCATCCATCCCC PP CCCC CCCC PP CCCCCCCCCC	-
GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	z
CAA' CCAA' N N N N N N N N N N N N N N N N N N	:
GTG CTA Y Y K K K K K K	Ŀ
AGC TCA TGC CAA N N N SAT	_
AAACCTA CTA Y Y AGG CTA CTA CTA CTA CTA CTA CTA	S
AGG CAG CAGG CAGG CAGG CAGG CAGG CAGG C	>
A A A B B B B B B B B B B B B B B B B B	⋖
GCCC AGG TGZ TGA' D D CCAA1	2
GAAGG GAGG GAGG GAGG GAGG GAGG GAGG GA	-
CCTG CCTG CCTG CCTG CCTG CCAC CCAC CCAC	2
CCA PF GTT FF C C C C C C C S S S S S S S S S S S	- .
C C C C C C C C C C C C C C C C C C C	-

FIG. 15e

1740	280	1800	009	1860	620	1920	640	1980	099	2040	089	2100
TG.		II		AC	_	TT	~	3AG	(-)	CT		၁၉
2	>	TA	_	3	Z	្ត្រ	111	Š	54	Ĭ		ຼວັ
S		IAA	Z	X.	Z	9	G	ည	¥	ξi	S	CT.
99	G	C¥.	ď	ŢĞ(ပ	T	Cz.	ည	¥	ర్ల	¥	1 6
Š	_	G		Ę		YC		AA		AT		.ye
ဗ္ဗ	<u>a</u>	AT	S	မ္ဌ	>	ა ა	➣	.09	Œ	ည	=	8
3GA	9	<u>ک</u>	ď	95	IJ	rg.	9	3CA	0	SGA	Œ	TGA
\ddot{S}	۵.	Š	~	AG	~	TG	ပ	IAG	24	SC.A.	O	50
ATG	æ	AAT	Z	299	G	TTC	رعا	735	¥	25	S	CAC
GAC	_	CTG	_	CAC	×	2	۵.	GAA	臼	gev	ප	959
AT		ည		ည		Ž	•	CA		TG		AT
ပ္ထ		$\tilde{\Sigma}$	O	5	0	¥	Like	¥.	-	Š		.y
999	S	AAT	-	ည	Ò	ည	¥	₹ 25	0	ည္တ	م	ATG
TT	_	IAA	M	IAT	Z	TG	3 =	ဦ	%	3GA	Œ	S
TAC	>	755	G	739	V	Č	=	AT(-	CAC	0	. L
3	_	ÄT		5		္တ	_	္ဌ	_	ပ္တ		22
2	-	;A6		5	0	9	-	ဗ္ဗ	ميليا	9	_	3A(
-23	=	<u> </u>	¥	/93	Œ	3	Œ	ဗ္ဗ	ၒ	ည	ð	£ 5.
ATTCTGTGCCGGGGACCCACGTGTACTTGGGCGATGACATGCCGGACCCCAGGGCTTGTG	I L C R G T H V · Y L G D D M P D P G L V	CTTGCAGGCACAAAGTGTGCAGATGGAAAATCTGCCTGAATCGTCAATGTCAAAATATT	LAGTKCADGKICLNRQCQNI	AGTGTCTTTGGGGTTCACGAGTGTGCAATGCAGTGCCACGGCAGAGGGGTGTGCAACAAC	SVFGVHECAMQCHGRGVCN	AGGAAGAACTGCCACTGCGAGGCCCACTGGGCACCTCCTCTGTGACAAGTTTGGCTFT	RKNCHC.EAHWAPPFCDKFGP	CAG	G G S T D S G P I R Q A E A R Q E A A E	TCCAACAGGGAGCGCCGGGGCCAGGAGCCCGTGGGATCGCAGGAGCATGCGTCTACT	SNRERGQGQEPVGSQEHAST	GCCTCACTGACACTCTAGCCCTCCCATGACATGGAGCCGTGACCAGTGCTGCTGC
99	5	AA(×	GTJ	^	Ç	Ħ	(GA)	Q	Š	≃	LCT
9		S		99		ည		2		3AG	(~)	CA
ဒ္ဌ	2	Ž	_ _	17	G	Σ	0	Š		9		791
GTG	ပ	AGG	ပ	CT1	Œ	GAA	Z	AAG	S	CAC	24	AC]
Ç	_	ည	~	ret.	>	3AA	×	991	ပ	CAA	z	S
ATJ	_	CT	1	AGI	S	AG(<u>ح</u>	GGAGGAAGCACAGACGGCCCCATCCGGCAAGCAGGAAGCAGGCAG	ပ	<u>ງ</u>	S	9 V

AGAGGAGGTCACGCGTCCCAAGGCCTCCTGTGACTGGCAGCATTGACTCTGTGGCTTTG	2160
CCATCGTTTCCATGACAACAGACACACACAGTTCTCGGGGCTCAGGAGGGGAAGTCCAG	2220
CCTACCAGGCACGTCTGCAGAAACAGTGCAAGGAAGGGCAGCGACTTCCTGGTTGAGCTT	2280
CTGCTAAAACATGGACATGCTTCAGTGCTGCTCCTGAGAGAGTAGCAGGTTACCACTCTG	2340
GCAGGCCCCAGCCTGCAGCAAGGAGGAAGAGCTCAAAAGTCTGGCCTTTCACTGAGC	2400
CCCCACAGCAGTGGGGGAGAAGCAAGGGTTGGGCCCAGTGTCCCCTTTCCCCAGTGACAC	2460
CTCAGCCTTGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTCTGATATGGCT	2520
TTTAGCATTTATTATGAAAATAGCAGGGTTTTAGTTTTAATTTATCAGAGCCCTGC	2580
CACCCATTCCATCTCCATCCAAGCAAACTGAATGGCATTGAAACAAAC	2640
TAGGAGAAAGGGCGGTGAACTCTGGCTCTTTGCTGTGGACATGCGTGACCAGCAGTACTC	2700
AGGTTTGAGGGTTTGCAGAAAGCCAGGGAACCCACAGAGTCACCAACCCTTCATTTAACA	2760
AGTAAGAATGTTAAAAAGTGAAAACAATGTAAGAGCCTAACTCCATCCCCGTGGCCATT	2820
ACTGCATAAAATAGAGTGCATCCCGCCC 2848	

 3
 3
 4
 3
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6
 6</t

			•			
S ICI	AAG	TIC	16T C	S &	ट्राट >	
GGG GAA GAG TGT GAT TGT GGA GAA GAG GAA TGT AAC AAC CCC TGC TGC AAT GCC TCT G E E C D C G E E E E C N N P C C N A S	13T	CTG TTG OCT CCT CRC CTG TCC CAG CAG CAG TCT CAC CTC CCC CAG TTC L L A P G T L C R B Q A R Q C D L P B F	ACC GOC AAG TCT COC CAC TOC CAC TTC TAC CAG ATG GAT GGT ACC COC TGT T G K S P H C P T N F Y Q M D G T P C	GAG GOC GOC CAG GOC TAC TAC TAC GOC ATO TOC CTC ACC TAC CAG GAG CAG TOC CAG B G G Q A Y C Y N G 14 C L T Y Q B Q C Q	CAG CTG TOG GGA CCC GCA CCT GCC CTC TCC TTC GAG AAA GTG AAT GTG Q L W G P G A R P A P D L C F E K V N V	
MAT	CAG O	g _	ACC T	9 0) -	
2 <u>2</u>	CAC CAG H Q	CTG TTG GCT CCT GCC CTG CCC CAG CCC AGG CAG TGT CAC CTC CCC	GGT ACC.	GAG CAG	AAg K	
2 <u>2</u>	700 700 C	کور م	GAT D	CAG OCC TAC TOC TAC AAC GOC ATG TOC CTC AOC TAC CAG Q A Y C Y N G M C L T Y Q	OCT GAC CTC TOC TTC GAG P D L C F E	
8 ~	ဥ္သ	ر د <u>ط</u>	ATG	TAC Y	13 F	
AAC	B [.] S	၁၁ တ	5 0	∯ [™]	ည္ဆ	% &
N AC	ည္သ	AGG ~	TAC	ည ၂	සු _	394
C EE	E E	29 √ .	71C F	ည္ဆ	S E	
æ æ	A A	S 0	AAC	ATC	5 2	ည္ဆင္ဆ
9 8	ट	S S S	ACC T	3 8	æ ∀	AAG K
E GA	GOS GOS GAG TIGT GCT CAC GOC TOC G A B C A H G S	೪ ≈	P . G	AAC	OCA CCT GOC R P A	OCA OGA GAC ACC TITI GGA AAC TGT GGA AAG GAC A A G D T F G N C G K D
E SA	SGC A	ဥ္သ	ည္သ	TAC Y	ই ~	15T C
E 5	8 5	<u>ඩු</u>	S ≖	ည္ည	gg _v	AAC
C 3	8 ~	T T	8 ~	TAC	8 5	ජි ය
GAT	AGG R	9	S	දි අ	8 ~	E
15T C	ر ا	لة 1	AAG K	S O	. 2	AC T
S 2	ACC T	A A	ည္တ	ည္တ	70C	250
E E	161 C	11 1	ACC	႘ၟ	ខ្លួ	త్ర ం
98	AAT	5 1	100	9 9	200	8 4

FIG. 17a

90	20	120	40	180	09	240	80	300	100	360	120	420	140
CGGAGCTGCCACTGGGCACCCCTTTCCCAAAGTGTTCAATGGATGCAACAGGAGGGAG	-	GAT	DRYLQSGGMCLSNUPDTRM	CTG	LYGGRRCGNGYLBDGEECDC	. ၁၁၅	GEEECNNPCCTLRP	GGGGGGGGAGTGTGCTCACGGCTCCTGCTGCACCAGTGTAAGCTGTTGGCTCCTGGGAC	—	CCTGTGCCGCGAGCAGGCCAGGCAGTGTGACCTCCCGGAGTTCTGTACGGGCAAGTCTCC	۵	CTA	HCPTNFYQMDGTPCEGGQAY
199	മ	CAG	8	.TG/	0	GA6	~	TGG	9	GTC	S)))	Y
GAG	8	CAC	\leftarrow	GTG	ပ	CCI	-1	TC	۵.	CAA	×	CCA	ð
CAG	~	AGA	9	AGA	Œ	TAC	⊣	၁၅၅	V	999	9	993	S
CAA	Z	၁၁၅	۵.	3GA.	妇	lTG	ပ ·	TT	-	Z Z	(355	ပ
ATG(ပ	CAT	=	<u> </u>	ပ	LAA.	z)CT	_	TG	ပ	rg A (22
<u>1</u> 66/	G	CAA(z	YGA.	Q	CIC	S	ra A (×	:TT	LCREQARQCDLPEFCTGKSP	TGT	ပ
AA	z	CTC	S	GA/	ম	ည	¥	:TG	ပ	Q V O	E	ည္တ	۵.
TT:	(2.	CTC	_	CTC	-1	:AA7	z	CAG	ď	22	م	VVCC	(—
GTO	>	TCI	ပ	TAT	>	.TG	ပ	CAC	×	CTC	-	CGT	5
AAA	×	ATG	Z	999	G	TGC	ပ	T 60	ပ	GAC	A	GAT	_
ည	۵.	66A	ပ	AAC	z	ဥ္ပ	۵.	TGC	ပ	TGT	ပ	ATG	*
TTT	(z.	GGT	<u>ن</u>	<u> </u>	ၒ	AAC	Z	JC	S	CAG	0	CAG	C)
ည	. م	GGT	ၒ	TGT	ر ن	AAC	z	ည်	ن ئ	AGG	22	TAC	>-
CAC	=	ICA	ις)	¥66	یہ	IGT	()	CAC		ည်		T.	f =.
366	(5	CAG	~ ~	993	~	3AA	ω 	CT	_	AG(~	NAC.	_
CT	د_	TG(. ,	ည		; A G(GT(``	AG(-:	733	-
733	_	ATC		GAO	٠.	AAG	<u> </u>	AGI		ည္ဟ	E4	CTA	
CTG	•	199	,	ATG		AAG	E	999	(2)	23	2	ည္တ	а,
GAG	₩ .	ACA	24	Tet	>	GAG	Œ	999	~	ToT	ပ	ACT	ပ
ည	9	99	Δ	GT	一 .	TG	S	99	5	ည		2	==

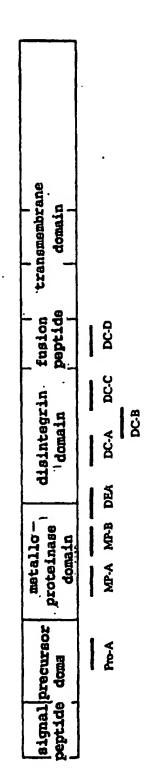
FIG. 17b

480	160	540	180	009	200	660	220	720	240	780	260	840	280
95	'n	95	()	ပ္				Ģ		ب	.	V	
္ပ္ဟ `		ľ.	_	5	O	ည်		ဋ္ဌ	œ	T.	3	T.G	(F)
AC.	7.	S	لعه	GA/	×	ပ္ပ	م	CI.	>-	¥	×	IAC	⊱
. 95. 95.	ح	K	\vdash	ည	V	Ţ	>	ET.	>	AC.		CA/	Œ
ŢĆ(=	GAC	0	3A.T	0	ည်		YY C	-	Ç.		'TT	
TG		SA S		ĞA	·	Š	_	ည္ထ	14	T.C	9	, L	Œ,
ည္ဟု -	- 3	YYC	S	S.A.	α	Č	Z	· 🕺	\vdash	GAC	⊱	CT	ÇZ.
ည	→	36	∢	:AT	==	T.C	S	99	9	AT	æ	ຼີ	S
S C	. · •	CI	>	AAC	2	GAC	മ	ဗ္ဗ	œ.	3TG		2	_
ည့ <i>့</i>		AT	_	ည		ŢĞ		ပ္ပ	_	ŢĆ		AC.	-
AGI,		ZS.	2.	VCI	0	္တ	-1	IGT	ပ	ည		Ş	Z
ည်	7	ပ္ပ	>	CA/	×	ပ္ပ	۵.	Z	0	AGG	်ပ	CAG	α
GGA T	4	3AA	×	.AG	2		~	AT	-	$\tilde{\mathcal{C}}$	۵.	.TG	ပ
S S S S	Y	ZY.	드	C. C.	Ħ	ည္တ	~	CAG	0	GAC	A	CAG	ď
LY ,	_	LIC	r-	, A A	(-)	AG	>	99		Ţ		ပ္ပ်	
ີວິ		ပ္ထ	ш	STO	ш	CTG	44	56	~	ည	7	99,	S
CA	₹ .	Ç	၁	ŢĞ	S	CŢ	S	ည်	ပ	CA1	Z	75 1	Œ
-	4		_	GAA	z	3AG	S	3AA	z	GA	9	CT	_
3TG	ز	<u>ک</u>	A	AT	=	Š	ð	AT(=	5	ပ	Ţ	ပ
AT(ည	<u>م</u>	GA(9	TG	ပ	ATC	_	GAG	GPEEEGDMLDPGLVMTGTKC	ATT	_
ည္တ	•	ပ္ပ	¥	AAG	×	CAG	ď	ATC	-	GAG	(-)	CAT	
AC .		5	^	Y.S.		T		CI		¥G.	_	AC	
AC.	-	ZY		STC	G	3		స్ట	H	ΥĞ	ल	CA	Z
CI.>	• 6	္ဌ	2	CŢ	ပ	GA.	×	CAC	[ຼ	<u>م</u>	CTA	>
CTGCTACAACGGCATGTGCCTCACCTACCAGGAGCAGTGCCAGCAGCTGTGGGGACCCGG	,	AGC	ARPAPDLCFEKVNVAGDTFG	AAA	N C G K D M N G E H R K C N M R D A K C	TGG	GKIQCQSSEARPLESNAVPI	TGACACCACTATCATGAATGGGAGGCAGATCCAGTGCCGGGGCACCCACGTCTACCG	0	AGGTCCTGAGGAGGGTGACATGCTGGACCCAGGGCTGGTGATGACTGGAACCAAGTG	ပ	TGGCTACAACCATATTTGCCTTGAGGGCAGTGCAGGAACACCCTCCTTCTTTGAAACTGA	G

FIG. 17c

AGGCTGTGGGAAGAGTGCCATGGGGTCTGTAACAACCAGCCGCCCTGGGCTGGGAACTGCCCTTGGGGTCTGTAACAACAACAACTGCCACTGGCTGCGGCCACGGGGCAGTATCGAAGGCTGCTGCGGCCACGGGGCAGTATCGAAGGCTGTGGCTGGTGGTGGTGGTGGTGGTGTTGGTGGCCATCTTGGGCTATGCTGGCTG	300	960 320	1020 340	.1080	1140 380	1183 394
AGGCTGTGGGAAG G C G K CCTGCCGGCTGG L P G W P G P W P GGTGCTGGCGTC V L A V K P S L K P S L K P S TTCTCAGAACAGC S Q N S G						
AGGCTGTGGG G C G CCTGCCGGGC L P G G P M G P M ACTCAAGCCC L K P L K P L K P S Q N S Q N	AAG/ K	TGGG) P P	GTCC V L	TCAG S A	AGCG S
AGGCTG G C CCTGCC L P G C G P G P G P L K L L K S Q	STGGGAA G K	3666CT6 6 ₩	TATGCC	GGCGGT A V	GCCCTC P S	GAACAG N S
	AGGCT(CCTGC(TGGGCC G P	6616C1 V L	ACTCAA L K	TTCTCA S Q

FIG. 18a Peptides used for the preparation of monoclonal antibody



 ${
m FIG.~18b}$ Peptide sequences used for the preparation of monoclonal antibody

No.	Name .	Sequence (N-terminal, C-terminal)
1	Pro-A	TTDSYKL VPAESMTN I C
2	MP-A	ADNREFQRQGKDLEKVKC
m	8-JR	FTRLHEFLDWRKIKC
4	P-2q	QLKPPGTACRGS SNSC
5	B-20	GTACRGSSNSCDLPEFC
9)-JQ	GKDSKSAFAKCELRDAKC
7	Q-2Q	QGGASRPV I GTNAVS I ETN I C
8	DE-A	LFNLPEVKQAFGGRKC

FIG. 19 Western blotting with anti-Meltrin monoclonal antibodies

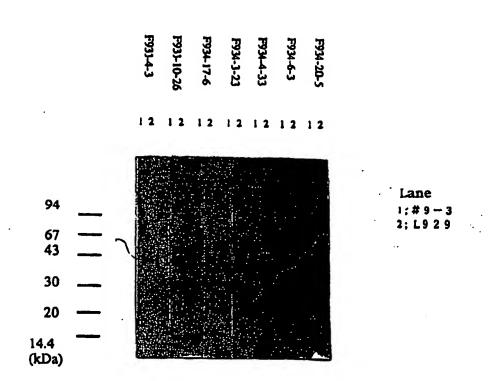
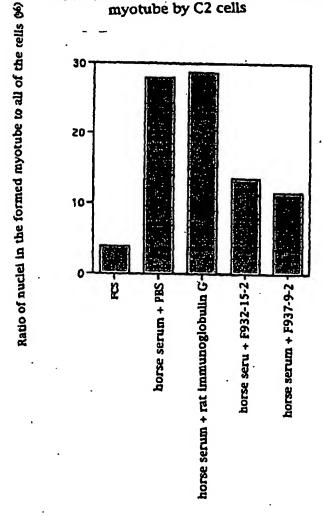


FIG.~20 Effects of anti-mouse Meltrin antibodies on the formation of myotube by C2 cells



Ratio of nuclei in the formed myotube to all of the cells (%)

FIG. 21 Effects of anti-mouse Meltrin antibodies on the formation of pit (bone-resorption area) in mouse unfractionated bone cells

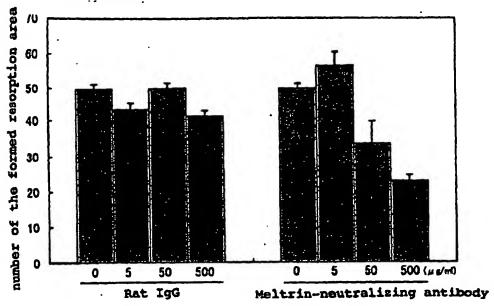


FIG. 22 Effects of anti-mouse Meltrin antibodies on the serum Ca values of the mouse fed with low Ca-content feed

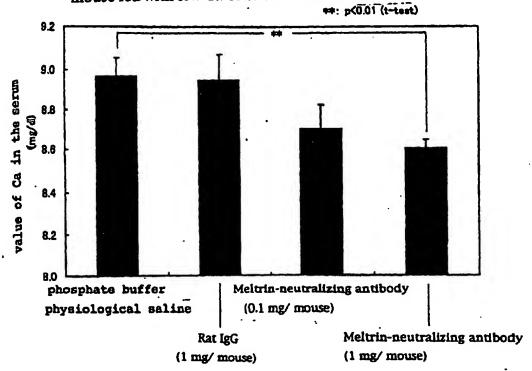


FIG. 23

8 8 **GCACAAAGTGTGCAGATGGAAAATCTGCCTGAATCGTCAATGTCAAAATATTAGTGTCT** ဟ z 0 ပ ď 2 z __ ပ -2 G 9 ¥ ပ **×**

120 40 TTGGGGTTCACGAGTGTGCAATGCAGTGCCACGGCGGGGGGTGTGCAACAACAAGA Z 4 2 Z Z ပ > ပ 24 Ç H O 0 × 4 ပ 四 Ħ ය

180 8 ACTGCCACTGCGAGGCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGCTTTGGAGGAA ى G Ľ, S (z, 8 <u>_</u> ပ ٦ ب ۵. ~ **B**e ¥ ¥ Œ ပ ပ

240 8 GCACAGACAGCGCCCATCCGGCAAGCAGATAACCAAGGTTTAACCATAGGAATTCTGG ပ [_ G ð Z A D 0 8 م ය S

300 100 IGACCATCCTGTGTCTTCTTGCTGCGCATTTGTGGTTTATCTCCAAAAGGAAGACCTTGA _ H × 24 7 _ > **>** > لتا ය ¥ 4 _ _ ပ _

360 120 S TACGACTGCTGTTTACAAATAAGAAGACCACCATTGAAAAACTAAGGTGTGTGCGCCCTT _ 2 > ပ 2 ._ M Œ H **-**74 × Z **-**بعا ٦.

140 CCCGGCCACCCATGGCTTCCAACCCTGTCAGGCTCACCTCGGCCACCTTGGAAAAGGCC ပ **~** G _ = ပ __ = ~ 0 ပ ۵. 0 لت ပ 2 م ہم 2

FIG. 23b

160 TGATGAGGAAGCCCCAGATTCCTACCCACACGAGGACAATCCCAGGAGATTGCTCCAGT 2 24 ۵. z 0 **×** م م **>**-S **A** ۵. ۵. **×** 24 æ

180 540 GTCAGAATGTTGACATCAGCAGCCCTCAACGCCCTGAATGTCCCTCAGCCCCAGTCAA S 0 م. 0 ۵. > Z د ය z u ۔ 2 S 0

909 200 ہم CTCAGGGAGTGCTTCCTCCCCACCGGGCTCCACGTGCACCTAGCGTCCCTGCCAGAC ~ ¥ ۵, ဟ ⋖ 2 ۵. ¥ 2 × ے م ۵.

CCCTGCCAGCCTGCACTTA 624

L P A K P A L 207

9	20	120	180	240	300	360 120	420 140
GCT	7	GAT	CTC	၁၁၄ -	3AC T	ည္ 🕳 .	TA Y
GGA	Œ	CAG R	rga D	3AG(R	ည်	rtci s) V V
GAG	~	CAC	ore.) L	ည	AAG K	CAG
CAG	≃	NGA D	IGA(ľACC T). A	၁၁	ည္သ
CAA	z	, 2007 1	GA/ E	TTG. C	itte L	7 7	၁၅၅
ATG(ပ	A T	၁၅၅	N N	icTo	TGT	GAG E
166,	ၒ	N	D D	S S	AAG K	TTC	TGT C
SAA.	Z	STC S	iGA/ E) V	TGI C	ÇAG E	ည
3TT	[2.	ict L	icte L	N N	CAG Q	9 20 4	ACC T
IGT(>). C	TAT Y) C	CAC	CTC	GGT
CAA	×	IAT(999		.TGC C	GAC	GAT D
22	۵.	1991	AAC N) }	.TGC C	TGT C	ATG M
TT	ഥ	ည်	999	AAC N	TCC S	CAG Q	CAG Q
\mathcal{Z}	۵.		761 C	AAC N	၁၁၁	AGG R	TAC Y
CAC	=	S S	AGG R	TGT	CAC	000 V	TTC
999	ပ	ÇA Ç	993. 8	GAA	GCT	O O	AAC
AC!	←	. ၂	9	GAG E	TGT C	GAG	ACC
))	¥	TA1 Y	.GGA	GAA E	GAG E	252 R	CCT
, 1907	G A A T G H P F P K V F N G C N R R E L	GACAGGTATCTGCAGTGGTGGAATGTGTCTCTCCAACATGCCAGACACCAGG	iTA1 Y	GAA	GCG A	TGC	16C
CGGAGCTGCCACTGGGCACCCCTTTCCCAAAGTGTTCAATGGATGCAACAGGAGGGAG	ర	GGACAGGTATCTGCAGTCGTGGAATGTGTCTCTCCAACATGCCAGACACCAGGAT D R Y L Q S G G G M C L S N M P D T R M	GTTGTATGGAGGCCGGAGGTGTGGGGAAGATGGGGAAGAGTGTGACTG L Y G G R R C G N G Y L E D G E E C D C	TGGAGAAGAAGAATGTAACAACCCCTGCTGCAATGCCTCAATTGTACCCTGAGGCC	GGGGGGGGAGTGTCCTCCTGCTGCCACCAGTGTAAGCTGTTGGCTCCTGGGAC	CCTGTGCCGCGAGCCAGCCAGTGTGACCTCCCGGAGTTCTGTACGGGCAAGTCTCC	CCACTGCCCTACCAACTTCTACCAGATGGTACCCCCTGTGAGGGGGGGCGAGGCCTA

FIG. 24b

480	160	540	180	900	200	099	220	720	240	780	260	840	980
CTGCTACAACGGCATGTGCCTCACCTACCAGGAGCAGTGCCAGCAGCTGTGGGGACCCGG	ن	AGCCCGACCTGCCCTCTGCTTCGAGAAGGTGAATGTGGCAGGAGACACCTTTGG	S	TG	၁	TV:	_	9	24	Te	ပ	¥5	(z
$\ddot{2}$	۵,	III	لعا	AAG	×	පු	۵.	TAC	>	AAG	M	ACT	_
799	S	ACC	-	93	¥	STS	>	ວຼາວ	>	ACC	<u>-</u>	SAA	(+ ¹
166		OVS	a	GAT	A	93	~	CAC	=	GGA	(5	ľTT(Cr.
CTG	_1	GGA	Ç	AGA	∞;	AAC	z	ACC	(ACT	<u> </u>	TC	·
CAG	0	SCA	¥	ATG	=	ට්ට	S	၁၅	` (5	ATG	- 	5	-
CAG	0	GTG	^	AAC	z	GAG	ய	993	<u>د</u>	376	_	SS	ری
TGC	CYNGMCLTYQEQCQQLYGPGPG	AAT	z	TGC	N C G K D M N G E H R K C N M R D A K C	TGGGAAGATCCAGTGTCAGAGCTCTGAGGCCCGGGCCCCTGGAGTCCAACGCGGTGCCCAT	ے	TGACACCACTATCATGAATGGGAGGCAGATCCAGTGCCGGGGCACCCACGTCTACCG	DTTIIINGRQIQCRGTHVYR	AGGTCCTGAGGAGGAGGTGACATGCTGGACCCAGGGCTGGTGATGACTGGAACCAAGTG	G P E E G D M L D P G L V M T G T K C	TGGCTACAACCATATTTGCCTTGAGGGGCAGTGCAGGAACACCTCCTTCTTTGAAACTGA	GYNHICLEGOCRNTSFFFF
CAG	0	GTG	^	AAG	_ 	ည	۵.	CAG	~	995		755	~
GAG	Œ	AAG	` ≥ d	AGG.	~ ~	990	~	VTC(_	ζ¥	^	735	
CAG		3AG	~ . (x)	CAC	_	ည္ဟ	ند	.AG/	_	.YCC	-	AGI	_
rac(<u> </u>	ŢŢ(fe.	3AA(~	;A60	~	25	G	TGG	Α .	26	0
CC	-	ည) ET		Ç		e GGA	02,	TGC	1	AGG	S
TC/		[2]	`	ATG	_	139	S	ATG	9	ACA	=	TTG	Œ
၁၁၅	-	ACC	_	TGA	Z	AGA	S	TGA	Z	GTG	Ω	ည်	
TGI	3	CTG	Ω	ACA	=	GTC	0	TCA	Ħ	A GG(ပ	ITT	ပ
CCA	æ	ည္ဟ	<u>а</u>	AGG	Q	AGT	ပ	ıÇ.	—	1991	Œ	TAT	-
ACG	S	CTG	V	GAA.	×	222	ď	T.Y.	-	991	臼	$\mathcal{Z}_{\mathcal{Z}}$	H
ACA	Z	GAC	م	STG	S	IGA.	-	ŠŠ	(-	TG/	压	CAA	Z
CCL	>-	Ö	~	CT	ပ	3GA/	×	CAC	(-)]C	a.	CT.	X
CI	ပ	AG	¥	AA	Z	TG	ت .	TGA	9	AGG	ပ	TGG	ပ

FIG. 24c

900	300	960	320	102	34(108(36(1140	38(1200	400	1960	420
AGGCTGTGGGAAGAGTGCAATGGCCATGGGGTCTGTAACAACAACCAGAACTGCCACTG	ပ	CAG	LPGWAPPFCNTPGHGGSIDS	CII		CCA	0	GT	Λ	S	a .	J.C	S
CCA	Ħ	CGA	9	CAT	-	AGG(ပ	CAGO	24	ည်	¥	ZY Z	=
CTG	ပ	TAT	-) J	¥	ACT.	_	CTT	Œ,	555	~)CT	
GAA	z	CAG	S	GGT	^	CAA	×	JCC	۵.	\TT(Œ.	CT	1
CCA	0	555	G	GTT	GPMPPESVGPVVAGVLVAIL	CAA	V L A V L M L M Y Y C C R Q N N K L G Q	lTG.	LKPSALPSKLRQQFSCPFRV	3GA/	SUNSGTGHANPTFKPEFRAP	CACC	HSPHHDKGHQFHGHTLLHS
CAA	Z	990	ဗ	AGT	>	ÇAA	Z	CAG	S	ည	۵.	CCAC	=
CAA	~	CCA	=	TGG	ပ	ACA	ď	GTT	(تد.	CAA	×	<u> </u>	ی
TAA	z	999	G	AGC	¥	CAG	24	ACA	0	TTT	Cz.	Y သွ	=
CTC	ပ	;ACC	ᇫ.	55	>	SCTG	ပ	igC _A	0	AAC	⊱	ATT	لعب
555	>	CAC	[.TG	A	CTG	ပ	CAG	~	ည	م	CCA	ð
\TG(G	CA	Z) JC	۵.	CTA	>	129	-1	CAA	Z	SC	==
7005	=	CTC	ပ	99	S	CT.	>	CAA	×	TGC	*	999	G
ATG	ဇ	CCTI	Œ	TCT	^	lga:1	=	TTC	S	TCA	×	CAA	×
3CA	Z	ည္ဟ	۵,	JEA(S	ເວຍ		ည	۵.	TGG	5	TGA	9
AGT(ပ	ည္ဟ	۵.	TG/	ഥ	[CA]	æ	Ž	-1	GAC		SCA	=
AGA,	×	366	Y	ၓ္တ	۵.	က္သ	-1) } } }	A	993	G	CCA	=
3GA	∠	3CT(25	۵.	55	>	SCTC	S	CAG	S	YCY	=
GTG(G	995	G	CTAT	32	099	¥	2291	۵.	GAA	Z	ည	۵,
CCT	ပ	TGC	۵.	ည	۵.	်ည	_	CA/	¥	TCA	O	CAG	S
AG	G	ပ္ပ	7	1 5	G	9	>	ACJ	-1	TT	S	ζ	×

TGGGGACGACCCGGATCCTCACTGAGCTGACCACAACAGCCACTACAACTGCAGCCACTG	132
G D D P D P H *	42
	•
GAICLACGCCACCCIGICCICCACCCCAGGACCACCIGGAICCICACAGAGCCGAGCA	138
CTATAGCCACCGTGATGGTGCCCACCGGTTCCACGGCCACCGCCTCCTCCACTCTGGGAA	144
CAGCTCACACCCCCAAAGTGGTGACCACCATGGCCACTATGCCCACAGGCCACTGCCTCCA	150
CGGTTCCCAGCTCGTCCACCGTGGGGACCACCCGCACCCTGCAGTGCTCCCCAGCAGCC	156
TGCCAACCTTCAGCGTGTCCACTGTCCTCCTCAGTCCTCACCACCCTGAGACCCACTG	162
GCTTCCCCAGCTCCCACTTCTCTACTCCCTGCTTCTGCAGGGCATTTGGACAGTTTTTCT	168
CGCCCGGGGAAGTCATCTACAATAAGACCGACCGAGCCGGCTGCCATTTCTACGCAGTGT	174
GCAATCAGCACTGTGACATTGACCGCTTCCAGGGCGCCTGTCCCACCTCCCCACCGCCAG	180
TGTCCTCCGCCCGCTGTCCTCGCCCTCCCCTGCCTGTGACAATGCCATCCCTC	186
TCCGGCAGGTGAATGAGACCTGGACCCTGGAGAACTGCACGGTGGCCAGGTGCGTGGGTG	192
ACAACCGTGTCGTCCTGCTGGACCCAAAGCCTGTGGCCAACGTCACCTGCGTGAACAAGC	198
ACCTGCCCATCAAAGTGTCGGACCCGAGCCAGCCTGTGACTTCCACTATGAGTGCGAGT	204
GCATCTGCAGCATGTGGGGGGGGCTCCCACTATTCCACCTTTGACGGCACCTCTTACACCT	210
TCCGGGGCAACTGCACCTATGTCCTCATGAGAGATCCATGCACGCTTTGGGAATCTCA	216
GCCTCTACCTGGACAACCACTACTGCACGGCCTCTGCCACTGCCGCTGCCGCCGCTGCC	222
CCCGCGCCCTCAGCATCCACTACAAGTCCATGGATATCGTCCTCACTGTCACTGGTGC	228
ATGGGAAGGAGGGCCTGATCCTGTTTGACCAAATTCCGGTGAGCAGCGGTTTCAGCA	234

FIG. 24e	
AGAACGGCGTGCTTGTGTGTGTGGGGACCACCACCATGCGTGTGGACATTCCTGCCC	2400
TGGGCGTGAGCGTCACCTTCAATGGCCAAGTCTTCCAGGCCCGGCTGCCCTACAGCCTCT	2460
TCCACAACAACACCGAGGGCCAGTGCGGCACCTGCACCAACAACCAGAGGGACGACTGTC	2520
TCCAGCGGGACGGAACCACTGCCGCCAGTTGCAAGGACATGGCCAAGACGTGGCTGGTCC	2580
CCGACAGCAGAAAGGATGGCTGCTGGCCCCGACTGGCACACCCCCCACTGCCAGCCCCG	2640
CAGCCCGGTGTCTAGCACCCCACCCCG 2669	

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

EP 96935358 A [0208]

4

Non-patent literature cited in the description

- KNUDSEN, K.A. et al. Expl. Cell Res., 1990, vol. 188, 175-184 [0003]
- MERGE, R.M. et al. J. Cell Sci., 1992, vol. 103, 897-906 [0003] [0003]
- DONALIES, M. et al. Proc. Natl. Acad. Sci., 1991, vol. 88, 8024-8028 [0003]
- ROSEN, G.D. et al. Cell, 1992, vol. 69, 1107-1119 [0003]
- MORRISON, T.G. Virus Res., 1988, vol. 10, 113-136
 [0005]
- BIOBEL, C.P. et al. Nature, 1992, vol. 356, 248-252 [0005]
- a guide to method and application. MICHAEL A.I. et al. PCR Protocols. Academic Press, 1990 [0035]
- SAMBROOK J. et al. Molecular Cloning a Laboratory Manual. Cold Spring Harbor Laboratory, 1989
 [0038]
- STANLEYT.; CROOKE; BERNALD LEBLEU. Antisense Research and Applications. CRC Publishing, 1993 [0044]

- BORREBAECK, C.A.K. J. Immunol., Meth., 1989, vol. 123, 157 [0064]
- TOSHIO KUDO. Tissue Culture, 1993, vol. 19, 61-65 [0064]
- CARTER et al. Pro. Nat. Acad. Sci, 1992, vol. 89, 4285 [0064]
- Molecular Cloning, a Laboratory Manual. Cold Spring Harbor Laboratory, 1989 [0069]
- SAMBROOK J. et al. Molecular Cloning, a Laboratory Manual. Cold Spring Harbor Laboratory, 1989 [0093]
- E. HARLOW; D.LANE et al. Antibodies, A Laboratory Manual. Cold Spring Harbor Laboratory [0093]
- CHOMCZYNSKI P.; SACCHI N. Anal. Biochem., 1987, vol. 162, 156-159 [0094]
- SMITH, D.B.; JOHNSON, K.S. Gene, 1988, vol. 67, 31-40 [0104]
- Molecular Cloning: A Laboratory Manual. Cold Spring Habor Laboratory, 1989 [0116]
- MIZUSHIMA S.; NAGATA S. Nucleic Acid Res., 1990, vol. 18, 5322 [0121]